DOCUMENT RESUME

ED 161 294 FL 009 820

TITLE	Report of the Phonology Laboratory, No. 2.
INSTITUTION	California Univ., Berkeley. Dept. of Linguistics.
PUB DATE	May 78
NOTE	119p.; One paper in this collection, "The Production of Tone," by John J. Ohala, is not included due to copyright restrictions.
AVAILABLE FROM	Editor, RPL, Phonology Laboratory, Department of Linguistics, University of California, Berkeley, California 94720 ($3.00)
EDRS PRICE	MF-$0.83 HC-$6.01 Plus Postage.
DESCRIPTORS	*Acoustic Phonetics; Articulation (Speech); Consonants; Distinctive Features; Intonation; Language Development; *Language Research; Language Universals; Neurolinguistics; Phonemics; *Phonetics; Phonological Units; *Phonology; Physiology; Research Reviews (Publications); *Speech; Stress (Phonology); Suprasegmentals; Syllables; Vowels

ABSTRACT

This is one of a series of reports intended to make the results of research available and to serve as progress reports. The following abstracts are included: (1) "Learning the Phonetic Cues to the Voiced-Voiceless Distinction: Preliminary Study of Four American English Speaking Children," Mel Greenlee; (2) "Learning the Phonetic Cues to the Voiced/Voiceless Distinction: A Comparison of Child and Adult Speech," M. Greenlee; (3) "Auditory Basis of Progressive Tone Spreading," H. R. Javkin; (4) "Towards a Model for Syntagmatic Processing," J. Kingston; (5) "Pseudo-explanations in Phonology," J.J. Ohala; (6) "Speculations on Pitch Regulation," J.J. Ohala; (7) "The Physiology of Stress," J.J. Ohala; (8) "Acoustic Reasons Why Labio-velars are Both Labials and Velars," J.J. Ohala and J. Lorentz; (9) "Control of Vocal Tract Length in Speech," C.J. Riordan; (10) "Vertical Larynx Position in Whispered Speech," C.J. Riordan; and (11) "Perceptual Correlates of Vowel Nasalization," J.T. Wright. The papers presented here are: (1) "Speech Aerodynamics and Phonological Universals," J.J. Jaeger; (2) "Towards a Phonetic Explanation for Universal Preferences in Implosives and Ejectives," H. Javkin; (3) "Phonological Notations as Models," J.J. Ohala; (4) "Southern Bantu vs. the World: The Case of Palatalization of Labials," J.J. Ohala; (5) "Explaining the Intrinsic Pitch of Vowels," J.J. Ohala and B.W. Eukel; (6) "Listeners' Ability to Identify Languages by Their Prosody," J.J. Ohala and J.B. Gilbert; (7) "The Story of (w): An Exercise in the Phonetic Explanation for Sound Patterns," J.J. Ohala and J. Lorentz; (8) "Testing Hypotheses Regarding the Psychological Manifestation of Morpheme Structure Constraints," J.J. Ohala and M. Ohala; and (9) "The Production of Tone," J.J. Ohala. The report concludes with announcements of available language tapes. (AMH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED161294

FL009720

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

John J. Ohala
Head, Phonology Lab

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

REPORT OF THE PHONOLOGY LABORATORY NO. 2 BERKELEY MAY 1978

REPORT OF THE PHONOLOGY LABORATORY NO. 2

UNIVERSITY OF CALIFORNIA, BERKELEY

MAY 1978

Table of Contents

Members of the Phonology Laboratory..... iv

Introduction..... v

Abstracts

Mel Greenlee: Learning the phonetic cues to the voiced-voiceless distinction: preliminary study of four American English speaking children..... 1

Mel Greenlee: Learning the phonetic cues to the voiced/voiceless distinction: a comparison of child and adult speech perception..... 2

Hector R. Javkin: Auditory basis of progressive tone spreading.... 4

John Kingston: Towards a model for syntagmatic processing..... 5

John J. Ohala: Pseudo-explanations in phonology..... 6

John J. Ohala: Speculations on pitch regulation..... 7

John J. Ohala: The physiology of stress..... 7

John J. Ohala and James Lorentz: Acoustic reasons why labio-velars are both labials and velars..... 8

Carol J. Riordan: Control of vocal tract length in speech..... 9

Carol J. Riordan: Vertical larynx position in whispered speech... 10

James T. Wright: Perceptual correlates of vowel nasalization..... 11

Papers

Jeri J. Jaeger: Speech aerodynamics and phonological universals.. 12

Hector Javkin: Towards a phonetic explanation for universal preferences in implosives and ejectives..... 31

John J. Ohala: Phonological notations as models..... 38

John J. Ohala: Southern Bantu vs. the world: the case of palatalization of labials..... 47

John J. Ohala: The production of tone..... 63

John J. Ohala and Brian W. Bukel: Explaining the intrinsic pitch of vowels..... 118

John J. Ohala and Judy B. Gilbert: Listeners' ability to identify languages by their prosody..... 126

John J. Ohala and James Lorentz: The story of [w]: an exercise in the phonetic explanation for sound patterns..... 133

John J. Ohala and Manjari Ohala: Testing hypotheses regarding the psychological manifestation of morpheme structure constraints. 156

Announcement..... 166

Cover: Hermann Grassmann (1809-1877), mathematician, philologist (discoverer of 'Grassmann's Law'), and pioneer in research on the acoustic structure of speech. See his work "Ueber die physikalische Natur der Sprachlaute" in Annalen der Physik und Chemie (1877) [Neue Folge] 1.606-629. For a review and an appreciation of his contribution to acoustic phonetics see: Svend Smith, "Analysis of vowel sounds by ear" in Archives Néerlandaises de Phonétique Expérimentale (1947) 20.78-96.

Errata, RPL No. 2

Page	Line (b= from bottom)	Incorrect	Correct
1	6b	-voice	+voice
1	5b	+voice	-voice
1	4b	differentiation	differentiator
2	16	+voice	±voice
5	12	presistent	persistent
5	5b	allows of	allows for
9	5	Psychology	Physiology
11	12	midnasalized	mid-nasalized
23	16	p d g z	b d g z
23	19	a'	q'
40	3b	were	was
42	11	inadequate for	inadequate to
66	This figure belongs on page 95.		
67	This figure belongs on page 66		
95	This figure belongs on page 67.		
121	4b	df=2.05	df=205
147	13	is 'underlying'	if 'underlying'

Members of the Phonology Laboratory

(July 1976 -- May 1978)

Susan Bowman	Teaching Assistant (to Dec. 1977)
William G. Ewan	Research Assistant (to June 1977) NIH Post-doctoral Fellow (July 1977 to June 1978)
Mel Greenlee	Teaching Assistant
Jeri J. Jaeger	Teaching Assistant, Research Assistant
Hector R. Javkin	Research Assistant (to Dec. 1977)
Sydney Karp	Technician (Oct. 1976 to June 1977)
Haruko Kawasaki	Teaching Assistant, Research Assistant
John Kingston	Secretary/Librarian (Oct. 1976 to Sept. 1977)
John J. Ohala	Professor of Linguistics, Director
Steve Pearson	Technician
Sandra Pinkerton	NIH Post-doctoral Fellow (from June 1977)
Carol J. Riordan	Research Fellow
Kenneth Whistler	Secretary/Librarian (from Oct. 1977)
James T. Wright	Research Assistant, Teaching Assistant

Introduction

The Report of the Phonology Laboratory is issued at irregular intervals in order to make the results of our research available to colleagues engaged in similar work and to serve as progress reports to the agencies which fund us. With the exception of reports on work in progress or short papers on instrumentation or software, the presentation of a paper here is not intended as a substitute for regular publication.

Copies of this report (and RPL 1) may be ordered from:

Editor, RPL
Phonology Laboratory
Department of Linguistics
University of California
Berkeley, CA 94720
(USA)

The cost of this volume is $3.00 (U.S.); the cost of RPL 1 is $2.50. Both prices include surface post. RPL is distributed free only in cases of agreed-upon exchange of reports.

The Phonology Laboratory is a facility of the Department of Linguistics, University of California, Berkeley. Its purpose is to provide space and facilities for instruction in phonetics and phonology and experimental linguistics in general and for advanced research by faculty and students. For support of research we gratefully acknowledge the following sponsors:

National Science Foundation (through grant no. BNS75-21417)

National Institute of Neurological and Communicative Disorders and Stroke, Department of Health, Education and Welfare (through institutional stipends accompanying the postdoctoral fellowships of Drs. Ewan and Pinkerton)

Committee on Research, UCB

Computer Center, UCB

Holbrook Fund, UCB

For support of instructional activities we thank the College of Letters and Science for generous assistance.

John J. Ohala

Abbreviations Used in References

AVPE	Archives Neerlandaises de Phonetique Experimentale
BLS	Proceedings of the Annual Meeting of the Berkeley Linguistics Society
BSTJ	Bell System Technical Journal
CLS	Proceedings of the Annual Meeting of the Chicago Linguistic Society.
JASA	Journal of the Acoustical Society of America
JSHR	Journal of Speech and Hearing Research
L & S	Language and Speech
Lg	Language
MIM	Monthly Internal Memorandum (Berkeley)
POLA	Project on Linguistic Analysis Reports
RPL	Report of the Phonology Laboratory (Berkeley)
SCRL	Speech Communication Research Laboratory (Santa Barbara)
STL-QPSR	Speech Transmission Laboratory (KTH, Stockholm), Quarterly Progress and Status Reports
WPP	Working Papers in Phonetics (UCLA)

Learning the Phonetic Cues to the Voiced-Voiceless Distinction: Preliminary Study of Four American English Speaking Children[1]

Mel Greenlee

Abstract

There has been considerable research into the phonetic cues for the phonological distinction [+ voice] for adult speakers of American English. It is well-established that for word-final obstruents, the duration of the preceding vowel is one of the most important, if not the most important cue for the differentiation of [+voice] and [-voice]. Other relevant cues may include characteristics of the obstruent release, presence or absence of voicing during the consonant closure, duration of obstruent closure, and ratio of vowel duration to consonant duration. Naeser (1970) found that the vowel duration difference is produced in a consistent way by American English speaking children even at the early age of 21 months, although what would be final voiced stops in adult speech were not consistently voiced by such young children; they often produce voiceless stops instead.

This paper reports the results of a small-scale study of 4 children aged 22-35 months which was carried out to determine how the various voicing cues found to be important in adult speech are acquired. Acoustic analyses of the children's production of selected CVC words was done using spectrographic and oscillographic displays. Some individual variation was found; nevertheless, the results, in agreement with Naeser, reveal that the vowel duration differences before [+voice] and [-voice] obstruents are a stable feature of the children's productions even though about half of the [+voice] stops were in fact voiceless. Closure duration of [-voice] stops were reliably shorter than [+voice] stops. The ratio of vowel duration to consonant duration was not as consistent a differentiation of the two obstruent categories.

---

1. Paper presented at the Seventh Annual Meeting of the California Linguistics Association, Fresno, Calif., April 30-May 1, 1977.

Learning the Phonetic Cues to the Voiced/Voiceless Distinction: A Comparison of Child and Adult Speech Perception[1]

'el Greenlee

In the field of child phonology it is commonly accepted that 'perception precedes production' of phonological distinctions in the adult language. However, as Barton (1976) pointed out, there has been little systematic study of children's speech perception during the early stages of meaningful speech production (i.e., from the second year of life up to about three years of age). Recently Zlatin and Koenigsknecht (1975) explored young children's perceptual capabilities with regard to the temporal acoustic cue Voice Onset Time, comparing children's perceptual identifications to those of adults.

The present study employed a comparative methodology similar to that used by Zlatin and Koenigsknecht, and to that used by Koenigsknecht (1968). The temporal acoustic cue investigated here was differential vowel duration, i.e., differences in vowel duration as a phonetic cue to the voiced-voiceless distinction in CVC minimal pairs such as back-bag. The object of this research was to determine whether young children, like adults, are capable of using vowel duration differences as the primary differentiator of the feature [+ voice] for obstruents in word-final position. Three-year-old children (N=12), six-year-old children (N=12), and adults (N=48), participated in two experiments, in which they were asked to identify (as voiced or voiceless) CVC words with uniformly voiceless final obstruents, but in which vowel duration was systematically varied. Children were also asked to identify a set of control stimuli, in which both closure voicing and vowel duration differences were present.

Test stimuli, unlike those of Zlatin and Koenigsknecht, were constructed from naturally spoken CVC's by computer splicing techniques. Each of eight CVC stimuli was presented with four vowel durations, in random order. Analysis of variance and Scheffe tests were used to compare the number of 'voiced' identifications given for each duration variant of a word across the three groups of subjects. For those CVC's which showed perceptual cross-overs in the voicing category of subjects' responses, cross-over points were compared across subject groups. In addition, the accuracy of children's identifications was compared for responses to the test stimuli versus the control stimuli.

Results of the two experiments indicate that both subject age and duration of the test stimuli significantly affected identification responses. Identification functions for the adult and six-year-old subjects show considerable similarity, while the three-year-olds form a group apart. Adults and six-year-olds evidence perceptual cross-over in their judgments for the test stimuli, but three-year-olds do not seem to change their identifications very much, regardless of variations in vowel duration. However, for both groups of children, the accuracy of identifications was greater for originally

---

1. Paper presented at the Tenth Annual Child Language Research Forum, Stanford University, Stanford, Calif., April 7-9, 1978.

voiced stimuli on the control set of words which contained more than one potential cue to the voicing distinction.

Two interpretations (at least) of these results are possible: (1) The identification task employed was too hard for the youngest children; thus negative results underestimate their true perceptual capabilities. (2) For three-year-olds, duration differences are not sufficient as a cue. Differences in identification on the test versus control words favor the second interpretation. Thus, there appears to be a complex relationship between developing production and speech perception which deserves further research. Children may consistently produce a phonetic difference (vowel duration) which they are unable to use as the sole indicator of a phonological contrast.

Auditory Basis of Progressive Tone Spreading [1]

Hector R. Javkin

Abstract

Relatively low frequencies (which include the F_0 of human speakers) are believed by many investigators to be processed by a different mechanism than most of the frequencies involved in speech segments. L. Hyman and R. Schuh ["Universals of tone rules", Linguistic Inquiry 5:81-115, 1974] noted that tone spreading, in which a linguistic tone spreads its domain in time, is always perseverative, never anticipatory. I hypothesized that the mechanism processing F_0 might take more time than the mechanism processing segments, creating an auditory image in which F_0 would seem to "lag behind" segments. The results of a psychoacoustic experiment involving 21 subjects support this: an F_0 change and a formant change were judged to occur simultaneously most frequently when the tone change occurred 10 msec before the center of the formant change. The apparent lag seems to cause listeners to perceive (and thus produce) tones as spreading into following segments. The explanation should be approached cautiously; although the perceptual facts seem clear, it nevertheless remains to be seen whether a 10-msec lag can lead to sound change.

---

1. Paper presented at the 92nd meeting of the Acoustical Society of America, San Diego, California, 15-19 November, 1976.

Towards a Model for Syntagmatic Processing[1]

John Kingston

Abstract

Confusion studies have demonstrated that the individual distinctive features of the speech signal differ in their contribution to the correct identification of segments under noise and filtering conditions. The results of one such study (Wang and Bilger 1973) were analyzed by means of a hierarchical clustering scheme (cf. Johnson 1967) to examine how much identification was dependent on the position of the consonant with respect to the vowel (i.e., CV vs. VC) and what the effect of symmetry vs. assymmetry in the response set was. Both parameters yielded positive results; in brief, initial position shows greater and more presistent differentiation than final position and a symmetric response set allows of a more coherent representation in terms of a feature description, in other words, a response bias. These results suggest that considerations such as the structure of the word as a syntagmatic (as opposed to paradigmatic) unit are at least as important to speech perception as taxonomic distinctions.

In: Proceedings of the 13th Regional Meeting, Chicago Linguistics Society, April 1977.

Pseudo-Explanations in Phonology[1]

John J. Ohala

Abstract

Explanations have the general form: 'x because y'. Ideally, 'y' should be distinct from 'x' (insofar as it represents an item in the accumulated knowledge of the world) but should be something previously known or something that people are ready to accept or believe. Pseudo-explanations, then, are those where 'y' is simply a cleverly (or not so cleverly) disguised form of 'x', i.e., where the 'explanation' is really of the form 'x because x'. Several examples of such pseudo-explanations exist in current phonology. These come in various forms: markedness or naturalness conventions, hierarchies, taxonomies, etc. Examples from the current literature will be given and analyzed logically. True explanations, that is, where facts and principles distinct from the thing being explained are referred to, are possible and will be given. Attention will be given to an aerodynamic model of speech, based on well-known laws of physics, which is capable of predicting the kinds of asymmetries encountered in the distribution of voiceless vowels, aspiration of consonants, and devoicing of consonants.

---

1. Paper presented at the Seventh Annual Meeting of the California Linguitics Association, Fresno, Calif., April 30-May 1, 1977.

Speculations on Pitch Regulation[1]

John J. Ohala

Abstract

This work briefly reviews some of the evidence for vertical larynx position being partly responsible for vocal pitch (F_0) variation.

The Physiology of Stress[2]

John J. Ohala

Abstract

This is a review of early and recent views on the physiological correlates of stress. Particular emphasis is given to the controversy regarding the relative contributions of the larynx and of the lungs in the implementation of stress. Original plethysmographic data (showing variation in lung volume during speech) are provided to show that the pulmonic system invariably participates only in the production of emphatic stress, not in ordinary non-emphatic stress. [120 references.]

---

1. Paper published in Phonetica 34:310-312 (1977).

2. Paper published in: Studies in Stress and Accent, ed. by L. M. Hyman. Southern California Occasional Papers in Linguistics, No. 4. Los Angeles, 1977. Pp. 145-168.

Acoustic Reasons Why Labio-Velars are Both Labials and Velars[1]

John J. Ohala

and

James Lorentz

Abstract

The phonological literature reveals an interesting asymmetry in the behavior of labio-velar speech sounds, i.e., w, u, kp, gb, etc., namely, that when nasalized or determining the place of articulation via assimilation of adjacent nasal consonants, they behave primarily like velars not labials, whereas when they become fricativized or determine the place of articulation of adjacent fricatives, they behave primarily like labials, not velars. The reasons for this will be explained by reference to well-known principles of acoustic phonetics. It will be shown that for nasals the most important of multiple oral constrictions is the rear-most constriction whereas in the case of fricatives it is the front-most constriction which is most important in determining the characteristics of the output sound. This case is offered as another example of the application of physics to phiology.

---

1. Paper presented at the 93rd Meeting of the Acoustical Society of America, University Park, Pennsylvania, 6-10 June 1977.

Control of Vocal Tract Length in Speech[1]

Carol J. Riordan

Abstract

Essential for the correct production of vowels is the accurate control of vocal-tract length. Perkell [Psychology of Speech Production (MIT, Cambridge, MA, 1969)] has suggested that two important determinants of vocal-tract length are vertical larynx position and lip spreading/protrusion, often acting together. The present study was designed to determine whether constraining lip spreading/protrusion induces compensatory vertical larynx displacements, particularly on rounded vowels. Upper lip and larynx movement were monitored photoelectrically while French and Mandarin native speakers produced the vowels /i,u,y/ first under normal-speech conditions and then with lip activity constrained. Significant differences were found in upper-lip protrusion and larynx position depending on the vowel uttered. Moreover, the generally low-larynx position of rounded vowels became even lower when lip protrusion was constrained. These results imply that compensatory articulations contribute to a contrast-preserving strategy in speech production.

---

1. Paper published in Journal of the Acoustical Society of America, Vol. 62, No. 4, October 1977, pp. 998-1002.

Vertical Larynx Position in Whispered Speech [1]

Carol J. Riordan

Abstract

A number of studies have shown that vertical larynx position varies depending on the vowel or consonant uttered. Voiceless stops, for example, are generally produced with a higher larynx than their voiced cognates. It is not clear, however, whether the larynx is perturbed upward for voiceless stops or downward for the voiced stops. The latter possibility is suggested by the increase in pharyngeal cavity size for voiced stop articulation, presumably to maintain the transglottal flow necessary for voicing [Bell-Berti, 1975], and by the similarity in larynx height data for sonorants and voiceless stops [Ewan, 1976]. The present experiment is designed to determine the effect on larynx height of sounds with aerodynamic requirements different from those of stops. Vertical larynx position is monitored photoelectrically while English speakers produce intervocalic /t,d,n,l,r,f,v,s,z/ first under normal speech conditions and then in whisper. Implications of these results are discussed.

---

1. Paper presented at the 93rd Meeting of the Acoustical Society of America, University Park, Pennsylvania, 6-10 June, 1977.

Perceptual Correlates of Vowel Nasalization[1]

James T. Wright

Abstract

The perceptual correlates of vowel nasalization were investigated by means of the multidimensional scaling of individual differences (INDSCAL). Steady-state vowels corresponding to [i, e, ɛ, æ, a, o, ʊ, u] and their nasalized counterparts were matched for tongue configuration across the oral and nasal conditions using optical palatography. Sixteen listeners heard all possible pairings of the stimuli and made similarity judgements on an interval scale. INDSCAL generated a unique analysis at three dimensions, which corresponded to the traditional vowel parameters high/low, front/back and nasal/non-nasal. Along the height dimension, high- and midnasalized vowels were consistently lower than their oral counterparts while the low-nasalized vowels were higher. The results provide evidence that listeners use a dimension corresponding to perceived degree of nasality in making judgements of similarity in the vowel space. They further suggest that perturbations in the first-formant region due to nasalization are perceived as changes in vowel height.

---

1. Paper presented at the 93rd Meeting of the Acoustical Society of America, University Park, Pennsylvania, 6-10 June, 1977.

Speech Aerodynamics and Phonological Universals [1]

Jeri J. Jaeger

I. Introduction

A. Explanations in Phonology

One of the most important current questions in linguistics is that of explanation; that is, what constitutes an explanation of certain observed linguistic data. In the area of phonological universals, there seem to be two types of methodologies: first, descriptions of phonological data which are intended as descriptions (for example, the work of Greenberg, Ferguson, and Hockett), and second, descriptions of phonological data which are intended as explanations (for example, the strength hierarchies and markedness conventions of Chomsky and Halle, Hooper, and Foley). These latter types of descriptions fail to be truly explanatory in that their primitives are defined in terms of the theory itself, rather than making reference to empirically verifiable principles outside the theory.

Recently some linguists, for example John Ohala (see Ohala 1977), have argued that those phonological processes which tend to be universal probably are so because they are caused by, and can therefore be explained with reference to, that which all speakers have in common: the human speech production and speech perception mechanisms. Most phonological patterns or sound changes which can be explained in terms of sociolinguistics, for example borrowings or fashion, will tend not to be universal. Since many aspects of human speech production and perception can be empirically investigated, explanations of phonological patterns based on these phonetic data can be truly explanatory, with primitives defined in terms of principles outside the theory, for example, principles of mathematics or physics. Further, theories based on empirically gathered phonetic data should be able to make predictions about phonological universals which can then be verified with reference to phonological data gathered from the world's languages. In this paper I will describe such a theory, present explanations of and predictions about several phonological universals, and then discuss these universals in detail with support from real language data.

B. A mathematical model of speech aerodynamics.

A computer-implemented mathematical model of speech aerodynamics has been reported in Ohala 1975a, 1976. This model basically works as follows: the user speci-

1. Paper to be published in Proceedings of the Fourth Annual Meeting of the Berkeley Linguistics Society, 1978, pp. 311-329.

fies starting values for various aspects of the modeled air passages, specifically, the volumes of the oral and pulmonic cavities, air masses in these cavities, and time-varying values for oral and glottal resistances and pulmonic force. The model has built into it various well-known principles of aerodynamics, for example, formulas for predicting when air flow will become turbulent, given its velocity and the dimensions of the chan nel it is passing through. The model is calibrated with reference to well-documented, measurable data about human speech aerodynamics. The model then outputs time-varying values for oral and glottal air flow and for air pressure in the oral and pulmonic cavities. Some of the values which the model outputs are compared with measurements taken from human speakers, to verify by extension the validity of the predicted values which are not easily measurable, such as pulmonic force or glottal air flow during obstruents.

The two main phonetic factors about which this model can make explanations are 1) voicing maintenance, and 2) introduction of turbulence into the speech air stream.

Voicing is a function of the pressure drop across the glottis; voicing can only be maintained when the air pressure above the glottis is less than that below the glottis, so that air can flow from the lungs across the glottis into the oral cavity, causing the vocal cords to vibrate. When there is a blockage of airflow in the oral cavity, whether it be a complete blockage as in stops, or a partial blockage as in fricatives or the narrow constriction for high vowels, the pressure of the air in the oral cavity will increase. If oral pressure becomes equal to subglottal pressure, voicing stops. So voicing maintenance depends on air pressure and velocity, oral cavity size, and the resistance of the oral constrictions, which are all specified or derived by the aerodynamic model.

Turbulence can be described as follows: when air is passing slowly through a tube or constriction, it moves in a smooth, or laminar flow; however, if air is forced to move through the constriction at faster and faster rates, at some point laminar flow will become turbulent, developing eddies in the air flow. A narrower constriction will cause turbulence to begin sooner, as air velocity increases. Turbulence, then, depends on the area of the constriction and air velocity, which are also specified or derived by the aerodynamic model.

These two factors, voicing maintenance and turbulence, are the key factors in the several phonological universals to be discussed here.

C. Methodology of this study.

Three predictions made by this model about phonological patternings related to aerodynamic factors were researched in the Stanford Phonology Archive. This Archive is a collection of computer-readable descriptions of the phonological systems of 221 languages (at the time of my research). Since the languages in the Archive have been chosen to represent as widespread and characteristic a sample as possible, no data from further languages was included, so that my sample would not be biased. Data was gathered on all languages which exhibited the phenomena in question, and simple counts were made to see whether or not the model's predictions were generally upheld. Examples and counterexamples were further researched to see whether something other than aerodynamic factors might be coming into play in any important way. Finally, languages were looked at in geographical groups, to make sure no purely areal features were surfacing; geographical locations of the languages discussed below are listed in Appendix 1.

II. Universals of Speech Aerodynamics: Three Examples

A. Voiceless vowels and vowel height.

The first prediction is also a well-known observed universal: in vowel systems in which only some vowels devoice, high vowels will devoice before non-high vowels. This universal was noted by Greenberg (1969) among others. The explanation is fairly simple: the narrower constriction for high vowels causes air pressure in the oral cavity to be greater than that during low vowels, so that devoicing due to supra and subglottal pressure equalization can occur sooner, as described above.

The data from the Archive supports this prediction, and reinforces earlier observations, as can be seen in Table 1. Out of 44 languages with voiceless vowels, 24 devoice only part of their vowel systems; of these 24, 20 either devoice only high vowels, as in Dafla, Greenlandic, and Guarani for example, or preferentially devoice high vowels, as in Japanese and Nyangumata. (Greenberg 1969 cites 10 additional languages not included in the Stanford Phonology Archive as having only high voiceless vowels. These languages are listed in Table 2.) Of the five languages which are exceptions (see Table 3), two, Morroccan Arabic and Paez, devoice only schwa. [1] In both of these languages, schwa is described as an overly short transitional sound; in the case of Paez, it is optionally inserted in certain environments, and is not phonemic. In both languages it

TABLE 1
Languages which only or preferentially devoice high vowels

Language	Devoices
1. Portuguese	i,u [-stress], in the environment of voiceless consonants.
2. Malayalam	i,y /e,a,o___#,C
3. Japanese	all, but high more often
4. Mandarin	following syllable in high-falling tone, syllable final high vowels under weak stress are voiceless following voiceless fricatives or aspirated fricatives.
5. Dafla	i,u / $\overset{C}{\circ}$___$\overset{C}{\circ}$
6. Akha	i / X___
7. Garo	i / s___n,r
8. Korean	y / medially, after fricatives, affricates, or aspir. cons.
9. Azerbaijani	I,ø,U,ɤ when short in unstressed open syllables
10. Gadsup	i / ___##
11. Western Desert	U in one morphological env.
12. Nengone	i,e,u,o (not ɛ,a,ɔ) / ___#
13. Nyangumata	all, but high more often, /__#
14. Greenlandic	i,u / $\overset{C}{\circ}$___$\overset{C}{\circ}$
15. Chipewayan	i, when second element of diphthong
16. Tunica	u, phrase-final preceeded by [k,hk] with penult. stress
17. Alabaman	i / #___s (free variation)
18. Mixtec	i / C___##
19. Campa	i / ___#
20. Guarani	i,ɨ,u / ___#

TABLE 2
Additional languages with only voiceless high vowels
listed in Greenberg 1969

Language	Devoices
1. Serbo-Croatian	i, u
2. Tadjik	i, u, ə
3. Awadhi	i, e, u
4. Uzbek	i, u
5. Dagur	i, e, u
6. Papago	all, high more extensively
7. Comanche	all but /a/
8. Shawnee	i̥ forces vowels in preceding syllable to devoice
9. Huichol	i, ʌ (+nigh), e
10. Chatino	i, u

TABLE 3
Languages whose only voiceless vowels are not high

Language	Devoices
1. Morocco Arabic	schwa / C̥___C̥
2. Paez	schwa / C̥___C̥
3. Hupa	[e,ə,o] but not [i,u]; the second half of long vowels devoice; only [e,ə,o] occur long.
4. Tarascan	[i,e,ɨ,a,u] not [o]; before pause when single, unstressed. /o/ is defective phoneme in this system.

devoices between voiceless consonants in unstressed syllables. Because it is a particularly short vowel and non-phonemic, its devoicing in this environment is not surprising; it is clear, however, that it is not a function of aerodynamic factors, but of some other phonetic factors.

The other two languages which devoice something other than high vowels also seem to do so for non-aerodynamic reasons. Hupa devoices /e,a,o/, but not its high vowels, which are described phonetically as: [i, I, ɨ, ɣ, ɯ]. However, vowel devoicing is an attribute of long vowels in this system, that is, the second half of long vowels devoice before pauses, and only /e,a/ and /o/ occur long. In fact, the high vowels in Hupa are sometimes analyzed as "lax" (short, unstressed) allophones of the phoneme /e/, so that they are phonemically never in a position to devoice.

Tarascan has a very odd system in that it devoices /i, e, ɨ, a/ and /u/, but not /o/. Vowels in this language devoice before pause juncture when occurring singly under weak stress, and /o/ does occur in this position, for example in the word Čénembo 'his house'. However, /o/ is reported to be a defective phoneme in Tarascan, of infrequent occurrence, which is never found in a number of environments, namely, after /w, th, t, tsh/ or /s/. One possible explanation for this imbalance is that /o/ may be a more recent addition to the vowel system than the others, and that the devoicing developed in the rest of the system before /o/ became part of the system; however, this is just a speculation. But it appears that none of these counterexamples seriously detracts from the original prediction, that high vowels are more likely to devoice than low vowels.

B. Effect of vowel height on obstruents.

The second prediction is that if a language has rules which either devoice, aspirate, fricate or affricate consonants before only some vowels but not others, it will be before high rather than low vowels. The reason is rather complex. As I have pointed out, the narrow constriction for high vowels causes increased oral pressure; this can delay voice onset time, as reported by Smith (1975). [2] The lag in VOT can be heard as being part of the preceding consonant, so that a voiced obstruent can be heard as voiceless, or a voiceless obstruent as aspirated. Further, high vowels are more likely to create turbulence than low vowels, because they offer a narrower channel for the air flow. This can affect the preceding obstruent in two ways: first, it can actually add frication to it; secondly, it is generally true that the frication on voiceless obstruent

segments is noisier and of higher amplitude than the frication on voiced segments, as noted in Stevens 1971; therefore it is possible that the presence of turbulence itself may favor a reinterpretation of preceding consonants as voiceless, aspirated or fricated.

Before I describe the Archive data I should point out that I did not include in my count cases described as 'palatalization' unless there was actual affrication involved; this is because palatalization is more a function of front-back assimilation (i.e. of tongue body) rather than an aerodynamic feature.

Languages seem to exhibit this phenomenon in two different ways, as shown in Table 4. Some have skewed phonological systems in which only the affricated allophone of an obstruent phoneme occurs before high vowels. Japanese is a well-known example of this situation, in that its alveolar stops and fricatives affricate before /i/ and /ɯ/. The Akan family shows a similar system, in which all alveolar and velar consonants affricate before high front vowels. Other languages have allophonic rules which devoice, affricate, aspirate or fricate single segments only in particular environments; the Gadsup language has three such rules, which are also listed in Table 4. Of the 23 languages, then, that exhibit the phenomenon in question, 20 do so before high vowels, and three before low vowels.

The counterexamples to this claim, listed in Table 5, are particularly interesting. In Yakut, a Turkic language of the USSR, voiceless velar stops are affricated before only low back vowels, and voiced velar stops are fricated before only low vowels, both front and back; the velar stops are also backed to uvulars. My only explanation for this is that in Yakut, as in various Turkic languages, there are front/back allophones of each obstruent (a function of 'palatalization' and vowel harmony), and that the back version of the velar phonemes are fricated, possibly due to the lesser agility of the tongue toward its root. However, this does not explain why even low <u>front</u> vowels cause frication of the voiced stop; I can think of no obvious phonetic explanation for this situation. On the surface, this is a counterexample to the predictions.

In Hupa the frication is used for emphasis, (i.e. t → tX in emphatic speech), which is clearly a function of perceptual rather than aerodynamic influences.

Lakhota, or Teton Dakota, has a very odd system in that it not only preferably affricates its aspirated stop phonemes before <u>low</u> vowels, but prefers to affricate them in the environment of <u>nasalized</u> vowels as well. The reason I find this strange is that there seems to be accumulating evidence that in general, nasals have a tendency to <u>block</u> aspiration, devoicing,

TABLE 4
Languages in which consonants devoice, aspirate, fricate or affricate in the environment of high vowels only

Language	Rules
1. Gbari	tʃ and dʒ occur only before /i,e/
2. Akan	d --> dz t --> ts g --> dǯ /___/i,e/ k --> tç h --> ç
3. Tigre	alveolar stops become fric. before /e/
4. Japanese	<table><tr><th></th><th>t</th><th>d</th><th>s</th><th>z</th><th>h</th></tr><tr><td>i</td><td>tʃ</td><td>dʒ</td><td>ʃ</td><td>dʒ~ʒ</td><td>ç</td></tr><tr><td>u</td><td>ts</td><td>dz</td><td>s</td><td>dz</td><td>ɸ</td></tr><tr><td>e,a,o</td><td>t</td><td>d</td><td>s</td><td>z</td><td>h</td></tr></table>
5. Ryukyuan	t --> ts, th --> tʃh, d --> dʒ /___j
6. Lanu	p,ph,b,m affric. before u u --->ɯ in this environment
7. Sa'ban	labials fricate / #__w
8. Selepet	ph --> ɸ,f / #___i, u
8. Gadsup	a) aspiration tends to be more frequent and pronounced before high vowels. b) β has less frication before low vowels. c) j → ǰ before high vowels.
10. Sentani	j --> dʒ / i, ɯ, j ____
11. Kunimaipa	l --> dɺ / high vowels
12. Nasioi	t --> ts / ?___i; t --> s / n,V___i
13. Nuangumata	p --> ɸ, pɸ / #___U
14. Greenlandic	t --> ts/ __i
15. Oneida	tʃ, dʒ occur only before [j,i]
16. Totonac	l --> ɬ / i___
17. Amahuaca	w --> β / env. of I
18. Carib	r --> dʒ / i,j ___V
19. Ticuna	dʒ --> j/ ___a
20. Jivaro	w --> β / ___i

TABLE 5
Languages in which consonants devoice, aspirate, fricate or affricate in the environment of non-high vowels only

1. Yakut: /k/ --> [qX] / ___ [a,ɔ]
   /g/ --> [ɣ] / ___ [a,e,ɔ,ɜ]

2. Hupa : /th/ --> [tX] in emphatic speech before /a/.

3. Lakhota: (Data from K. Whistler and R. Van Valin)

	/ph/	/th/	/kh/
i	h	h	h
e	h~X	h	h
u	h	h~X	h
ɔ	X	X	h~X
ɜ	X	X	X
ĩ	h	h	X
ɔ̃	X	X	X
ɜ̃	X	X	X

thipi	[tʰipi]	'house'
ophethū	[opʰetXū or opXetXū]	'to buy'
thāka	[tXāka]	'big'

or affrication of preceding segments [3], so that it seems that a language which preferentially affricates its obstruent phonemes in nasal environments is unusual, if not aberrant. However, the explanation for this phenomenon may in fact lie in its very strangeness. Consider the following speculation: Lakhota has both a voiceless aspirated and a voiceless unaspirated stop series. In certain environments, however, the aspiration tends to be lessened or damped, specifically around low vowels for aerodynamic reasons, and around nasals, probably for perceptual and aerodynamic reasons, since air escaping from the nose has less turbulence and therefore less perceptible noise than air escaping from the mouth. Therefore, in order to preserve the perceptual distinction between aspirated and unaspirated voiceless stops, frication is added to the segments in just those environments where it does not occur naturally. Because this frication occurs in the environment of low and back vowels, it takes on the same tongue position and occurs as velar frication. While this solution is merely speculation, it seems to be plausible, and if it were true, it would in a sense reconfirm the prediction that the more likely place to find devoicing, aspiration, frication, or affrication of obstruents is before high rather than low vowels. [4]

C. Geminates and Long Consonants

The third prediction is as follows: If a language has both voiced and voiceless obstruents, but geminates ony part of its obstruent system, it will have long voiceless rather than long voiced obstruents.

The explanation is that a stop closure of long duration will allow air pressure in the oral cavity enough time to equalize with sub-glottal pressure and cause voicing to stop; this is also true of the narrow constriction for fricatives, but not true of nasals, which of course allow air to escape from the nose, disallowing oral pressure build-up. Voiced geminate obstruents are more unstable than voiceless; over time they should either devoice, or devise some means of prolonging voicing, for example, becoming prenasalized or imploded. While this is clearly a diachronic prediction, a synchronic manifestation would be languages with geminate systems skewed in a voiceless direction, or with geminate or long obstruents only among the voiceless series.

In making this survey I of course discounted languages which have no voiced obstruent segments as phonemes or allophones. However, I did include languages which lack phonemic voicing but do have voiced allophones contrasting with the geminate obstruents. Exam-

ples of this are Ojibwa and Delaware.

The data from the Archive, displayed in Table 6 again supports our prediction: of 20 languages with skewed geminate systems, 18 either have only long voiceless obstruents or have systems skewed in a voiceless direction, not counting nasals. Japanese is again the classic case which exhibits this phenomenon in its phonological system, as it has both voiced and voiceless obstruents, but only geminate voiceless. Lak, a Caucasian language, is a further example. Awiya, an African language spoken in Ethiopia, demonstrates a voiceless skew, with two voiced and two voiceless geminate stops, but only voiceless geminate fricatives and affricates.

One of the counterexamples (see Table 7) appears to be in fact a verification of the claim that voiced geminates are diachronically unstable. Armstrong (1934) analyses Somali as having voiceless obstruent phonemes of rather limited occurrence, and voiced obstruent phonemes with a number of allophonic variations. Intervocalically, voiced obstruent phonemes can occur either short, in which case they are fully voiced and somewhat spirantized, or long, in which case they are described as "[not] sound[ing] fully voiced", or having "weak voicing", or none at all. What seems to be happening is that these long voiced segments are in the process of devoicing, as our model predicts they are likely to do.

The second counterexample, that of Island Carib, seems to be an example of a skew due at least in part to systemic reasons; that is, the voiceless segments /p/ and /k/ only occur initially, and long consonants only occur finally, so that /p/ and /k/ are never in a position to occur long. (p/b and k/g neutralize medially, so that the phonemicization of this language is problematic.) However, of the two voiceless stops which do occur finally, /t/ does occur long, but /c/ does not, and all the voiced obstruent phonemes do occur long. So while in one sense this stands as a counterexample, it appears that whatever is causing the skewing toward voiced geminates is systemic, a function of the somewhat asymmetrical phonological pattern, rather than due to aerodynamic factors. That is, non-phonetic factors have overridden the phonetic factors in determining the behavior of the sounds in this case.

It is interesting to note in passing that the most frequent type of geminate or long consonant in the world's languages is clearly the nasal geminate. This fact was briefly mentioned by Ferguson (1975), but no explanation was offered. As indicated in Table 9, the Archive shows 72 languages with some type of long geminate consonants, 67 of which have long nasals. This leaves only five languages which have only non-nasal long consonants. There are sixteen languages whose only

TABLE 6

Languages which either have only long voiceless obstruents, or geminate systems skewed in a voiceless direction.

Language	Language
1. Finnish	10. Lak
2. Icelandic	11. Yurak
3. Walamo	12. Evenki
4. Awiya	13. Greenlandic
5. Kanuri	14. Alask. Eskimo
6. Malayalam	15. Ojibwa
7. Kurux	16. Delaware
8. Japanese	17. Karok
9. Maranungku	18. Alabaman

EXAMPLES

Japanese

p	d	g	z
p	t	k	s
p:	t:	k:	s:

Lak

b	d			g	
p	t	ts	ǯ	k	q
p:	t:	ts:	ǯ:	k:	q:
p'	t'	ts'	ǯ'	k'	a'

Awiya (skewed system)

b		ɖ	g	dz	dʒ	z	ʒ
p	t		k	ts	tʃ	s	ʃ
p:	t:	ɖ:	g:	ts:	tʃ:	s:	ʃ:

TABLE 7

Languages which either have only long voiced obstruents, or have a skew towards voiced geminates.

1. Somali: all geminates are phonemically voiced, but are in the process of devoicing.

2. Island Carib:

b	d		g	f	s	h
b:	d:		g:	f:	s:	
p	t	c	k			
	t:					

/p/ and /k/ only occur initially; long variants only occur finally.

TABLE 8
Geminate Nasals

Number of languages with geminate or long consonants of any type.....72

Number of languages with geminate or long nasals.....67

Number of languages whose only long or geminate consonants are nasals.....16

Number of languages whose only long or geminate consonants are nasals and liquids....9

Languages without long nasals, which have other long consonants:

Language	Has Long:
1. Atayal	t, l
2. Cheremis	r
3. Iraqw	s
4. Totonac	p,t,k,q,ɬ
5. Tarascan	k

long consonants are nasals, and another nine which have only long nasals and liquids; combined, these two categories comprise over a third of the languages in this count. The interesting question is this: why are long nasals more frequent than long voiceless oral obstruents; is this just a function of the greater stability of nasals over time, as noted by Ferguson 1975, and Ohala 1975b, or are there other factors, perhaps aerodynamic, involved? The answer will have to be the topic of some future research.

III. Conclusion

In his recent dissertation Hector Javkin has argued that most sound changes of the universal sort can be explained with reference to either primarily articulatory, acoustic, or perceptual facts, and that each of these areas will be able to explain a different type of sound change. Because synchronic patterns are nothing more than the output of previous sound changes, as Greenberg 1966 argues, it follows that phonological universals should also be explainable in these terms. I have shown in this paper that certain facts about the aerodynamic aspects of speech production can be used to predict and explain a particular body of phonological data. The mathematical model of speech aerodynamics discussed here does not pretend to be a theory for explaining all possible speech patterns, but rather only some which are directly a function of aerodynamic facts. Because phonological universals obviously are caused by a wide range of phonetic factors, articulatory, acoustic and perceptual, any theory which hopes to be able to explain fully all phonological universals will have to include data from all of these areas; and since our knowledge in many facets of these fields is at this point sketchy at best, much more research will need to be done before such a comprehensive theory can be developed. However, it is clear that the basis for truly explanatory theories about phonological universals will have to be empirically gathered phonetic data, such as I have described in this paper.

ACKNOWLEDGMENTS

I would like to thank Johanna Nichols, Karl Zimmer, and John Kingston for useful comments on this paper. I would especially like to thank John Ohala for his thoughtful guidance and criticism throughout this project. I would also like to express my appreciation to the numerous people who helped me with the language data. This paper was funded in part by an NSF grant to the U. C. Berkeley Phonology Laboratory, and by a University of California Regents' Fellowship.

FOOTNOTES

1. I originally included Puget Sound Salish as a language with voiceless schwa, but have since had several Salishanists explain to me that this is incorrect.

2. Smith's results are somewhat complex in that he finds that fully voiced stops will have more pre-voicing before high than low vowels, while voiceless stops will have a much longer VOT before high than low vowels.

3. The incompatibility of nasalization with oral stops and fricatives is discussed in Ohala 1975a, and with frication on vowels and glides in Ohala 1977. The reason cited is the necessity of high oral pressure for noise bursts in the former case, and turbulence in the latter; the release of air through the nose makes high oral pressure difficult to accomplish. In going through the Archive data I came across a large number of languages with rules which, for example, devoice or aspirate obstruent phonemes everywhere, except in the environment of nasals, where they are voiced or unaspirated. However, there is some evidence for an affinity between /h, ʔ/ and nasality (see Matisoff 1975), and 'spontaneous nasalization' is attested around /s/ and /h/ (See Ohala 1975b). This entire area is very unclear and awaits further investigation.

4. This is not a totally unprecedented type of argumentation. Jakobson (1962) has used a similar argument to explain why Ukranian has palatal consonants only before low back vowels.

APPENDIX 1
Geographical location of languages discussed

Language	Area	Language	Area
Finnish	Europe	Gbari	Africa
Portuguese	"	Akan	"
Icelandic	"	Tigre	"
Serbo Croatian	"	Walamo	"
Moroccan Arabic	Morocco	Awiya	"
Awadhi	India	Kanuri	"
Malayalam	"	Somali	"
Kurux	"	Iraqw	"
Tadjik	USSR	Greenlandic	N. America
Uzbek	"	Alask. Eskimo	"
Yakut	"	Chipewayan	"
Lak	"	Tunica	"
Cheremis	"	Alabaman	"
Azerbaijani	"	Papago	"
Yurak	"	Comanche	"
Evenki	"	Shawnee	"
Dagur	East Asia	Hupa	"
Mandarin	(Peking)	Oneida	"
Atayal	"	Lakhota	"
Japanese	"	Ojibwa	"
Ryukyuan	"	Delaware	"
Lahu	S.E. Asia	Karok	"
Dafla	"	Mixtec	Mexico
Akha	"	Huichol	"
Garo	"	Chatino	"
Korean	"	Tarascan	"
Gadsup	Pacific	Totonac	"
Sa'ban	"	Campa	S. America
Selepet	"	Guarani	"
Sentani	"	Paez	"
Kunimaipa	"	Amahuaca	"
Nasioi	"	Carib	"
Nyangumata	Australia	Ticuna	"
West. Desert	"	Jivaro	"
Nengone	"	Island Carib	"
Naranungku	"		

REFERENCES

Abbreviations used:

IJAL = International Journal of American Linguistics
JASA = Journal of the Acoustical Society of America
UCPL = University of California Publications in Linguistics

Armstrong, Lilias E. [1934] 1964. The Phonetic Structure of Somali. Ridgewood, New Jersey: Gregg Press.

Aschmann, H. P. 1946. Totonaco phonemes. IJAL 12:34-43.

Bloomfield, L. 1956. Eastern Ojibwa. Ann Arbor: Univ. Michigan Press.

Bright, W. 1957. The Karok Language. Berkeley: U. C. Press. (UCPL 13).

Chao, Y. R. 1968. A Grammar of Spoken Chinese. Berkeley: U. C. Press.

Chomsky, N. & M. Halle. 1968. The Sound Pattern of English. New York: Harper & Row.

Clayre, I. 1973. The phonemes of Sa'ban: a language of Highland Borneo. Linguistics 100:26-46.

Dirks, S. 1953. Campa (Arawak) Phonemes. IJAL 19:302-304.

Douglass, W.H. 1964. An Introduction to the Western Desert Language. Oceania Ling. Monograph 4.

Einarsson, S. 1945. Icelandic Grammar. Baltimore: John Hopkins Press.

Ferguson, C.A. 1963. Assumptions about nasals: A sample study in phonological universals. In: J. Greenberg, (ed), Universals of Language. MIT Press, pp. 53-60.

_____ 1975. Universal tendencies and 'normal' nasality. In C. A. Ferguson, L.M. Hyman & J. J. Ohala (eds), Nasalfest. Stanford: Language Universals Project, pp. 175-196.

Foley, J. 1970. Phonological distinctive features. Folia Linguistica 4:87-92.

_____ 1975. Nasalization as a universal phonological process. In C. A. Ferguson, L. M. Hyman & J. J. Ohala (eds). Nasalfest. Stanford: Language Universals Project, pp. 197-212.

Foster, M. L. 1969. The Tarascan Language. Berkeley: U. C. Press. (UCPL 56)

Frantz, C. I. & M. E. Frantz. 1966. Gadsup phoneme and toneme units. Pacific Linguistics, Series A, 7:1-11.

Garvin, P. L. 1950. Wichita I: Phonemics. IJAL 16:179-84.

Gerdel, F. 1973. Paez Phonemics. Linguistics 104:28-48.

Golla, V. K. 1970. Hupa Grammar. Univ. Calif. Berkeley Doctoral Dissertation, unpub.

Greenberg, J. 1966. Synchronic and diachronic universals in phonology. Language 42:508-17.

_____ 1969. Some methods of dynamic comparison in linguistics. In J. Puhvel (ed). Substance and Structure of Language. UCLA: Center for Research in Languages and Linguistics, pp.147-204.

_____ 1970. Some generalizations concerning glottalic consonants, especially implosives. IJAL 36:123-145.

Haas, M. R. 1940. Tunica. Handbook of American Indian Languages, V. IV, New York: J. J. Augustine.

Hockett, C. F. 1963. The problem of universals in language. In J. Greenberg (ed). Universals of Language. MIT Press, pp. 1-29.

Hooper, J. B. 1976. An Introduction to Natural Generative Phonology. New York: Academic Press.

Jakobson, R. 1962. Remarques sur l'évolution phonologique du Russe comparée a celle des autres langues Slaves. In Selected Writings, I: Phonological Studies. 'S-Gravenhage: Mouton & Co., pp. 7-116.

Javkin, H. R. 1977. Phonetic Universals and Phonological Change. U. C. Berkeley, doctoral disseration, unpubl.

Kaufman, T. 1971. Tzeltal phonology and morphology. Berkeley: U. C. Press (UCPL 61).

Krueger, J. R. 1962. Yakut Manual. Indiana Univ. Publ., Uralic & Altaic Series, V. 21. Bloomington: Indiana University Press.

Lounsbury, F. 1953. Oneida Verb Morphology. Yale University Press.

Lukas, J. 1937. A study of the Kanuri Language. London: Oxford University Press.

Matisoff, J.A. 1975. Rhinoglottophilia: the mysterious connection between nasality and glottality. In C. A. Ferguson, L. M. Hyman & J. J. Ohala (eds). Nasalfest. Stanford: Language Universals Project, pp. 265-288.

McAlpin, D. W. 1975. The morphophonology of the Malayalam noun. In Schiffman & Eastman (eds), Dravidian Phonological Systems, Seattle: U. Washington Press.

McElhanon, K. A. 1970. Selepet phonology. Pacific Linguistics, Series B, no. 14.

Minn, E. K. 1956. Studies in Cheremis. Indiana Univ. Publ., Slavic and East European Series, Research Center in Anthropology, Folklore & Linguistics. Bloomington: Indiana University Press.

Newman, S. 1956. Zuni Grammar. Univ. of New Mexico,

Publications in Anthropology 14.

Ohala, J. J. 1975a. A mathematical model of Speech aerodynamics. In: Speech Communication. Proceedings of the Speech Communication Seminar, Stockholm, Aug 1-3, 1974. Vol. 2 (Speech Production & Synthesis by Rule) pp. 65-72.

_____ 1975b. Phonetic explanations for nasal sound patterns. In C. A. Ferguson, L. M. Hyman & J. J. Ohala (eds). Nasalfest. Stanford: Language Universals project, pp. 289-316.

_____ 1976. A model of speech aerodynamics. Report of the Phonology Laboratory, Berkeley, 1:93-107.

_____ 1977. Phonological notations as models. Paper presented at 12th Intl. Congress of Linguists, 29 Aug- 2 Sept. 1977, Vienna, Austria.

Palmer, F. R. 1962. The Morphology of the Tigre Noun. London: Oxford University Press.

Pence, A. 1966 Kunimaipa phonology: hierarchical levels. Pacific Linguistics, Series A, 7:49-67.

Pfeiffer, M. 1972. Elements of Kurux Historical Phonology. Leiden: E. J. Brill.

Rand, E. 1968. The structural phonology of Alabaman, A Muskogean language. IJAL 34:94-103.

Ristinen, E. K. 1965. On the Phonemes of Nenets. Ural-Altaische Jahrbücher 36:154-164.

Schachter, P. & V. Fromkin. 1968. A Phonology of Akan: Akuapem, Asante and Fante. UCLA Working Papers in Phonetics, no. 9.

Shipley, W. F. 1964. Maidu Grammar. U. C. Press, Berkeley.(UCPL 41)

Smith, B. 1975. Effects of vocalic context, place of articulation, and speaker's sex on "voiced" stop consonant production. JASA 58, Suppl. 1, p. 561 (abstract).

Stevens, K. N. 1971. Airflow and turbulence noise for fricative and stop consonants: static considerations. JASA 50:1180-1192.

Taylor, D. 1955. Phonemes of the Hopkins dialect of Island Carib. IJAL 21:233-241.

Voeglin, C. F. 1946. Delaware, an Eastern Algonquian Language. Linguistic Structures of Native America, pp. 130-157. (Viking Fund Pub. in Anthro. no. 6) New York: Wenner-Gren Foundation.

Whiteley, W. H. 1958. A Short Description of Item Categories in Iraqw. Kampala: East African Institute of Social Research.

TOWARDS A PHONETIC EXPLANATION FOR UNIVERSAL PREFERENCES IN IMPLOSIVES AND EJECTIVES[1]

Hector Javkin

Glottalic consonants are produced by closing the glottis and using movements of the oral cavity, particularly the upward or downward movements of the larynx, to compress or rarefy the air in the oral cavity. Implosives involve a downward movement of the larynx and produce a relatively low pressure; ejectives involve an upward movement and produce a compression of the air in the oral cavity. Haudricourt, in 1950, first noted that implosives and ejectives show strongly opposite tendencies for place of articulation. Languages show a preference for ejectives in the order: velar, alveolar, labial, while implosives occur most often in the opposite order. Greenberg, in an extensive work in 1970, firmly established these tendencies and put them into implicational relations. He found that a language will only have velar implosives if it has alveolar and labial implosives; it will only have alveolar implosives if it also has labial ones. For ejectives, the opposite implications hold: a language will only have labial ejectives if it has alveolar and velar; it will only have alveolar if it has velar.

Despite some counter-examples found by Campbell (1973) in some Mayan languages, Greenberg's observations seem to be correct. The tendencies were generally confirmed in a count conducted by the Stanford Phonology Archive, which appears below.

Stops	Labial	Alveolar	Palatal	Velar	Uvular
Implosive	17	15	2	2	0
Ejective	26	29	7	31	15

It should be noted that palatals and uvulars do not maintain these tendencies, since these places of articulation tend to disfavor stops. Secondly, although the numerical preferences for ejectives at the three major places of articulation do not seem overwhelming, the implicational relations still hold in the languages in the archive, thus upholding Greenberg's claim.

The fact that this tendency holds in so many languages suggests a phonetic basis. It is not surprising, then, that phonetic explanations have been suggested for this tendency. What is surprising is that there has been relatively little phonetic data on these consonants, and that so much remains to be done in finding the correct explanation.

Wang (1968:8) suggests the following:

---

1. Paper published in Proceedings of the Third Annual Meeting of the Berkeley Linguistics Society, 1977, pp. 559-565.

"...there seems to be a tendency for compression to occur with smaller volumes of air, while rarefaction occurs with larger volumes. This may account for the fact that labial ejectives are comparatively rare, and that velar implosives have not been reported."

Greenberg, in a footnote, acknowledges benefitting greatly from a discussion with Wang concerning the phonetic basis of some of the tendencies involving these sounds. A few paragraphs later, he says (1970:139):

"The point of articulation hierarchies of ejectives and injectives are obviously based on preference for a small and large air chamber, respectively. It is also clear that with the same thoracic pressure it is easier to build up compression in a smaller chamber. Injectives are usually voiced and involve leakage of air through the descending cords; such leakage is more easily tolerated from a large chamber."

Finally, Ladefoged (1971:43) cites Greenberg's explanation as the basis for these tendencies.

I shall ignore the confusion evidenced by the last line quoted from Greenberg; it is obvious that, given that implosives rarefy the air in the oral cavity, air will leak into the cavity, not from it. As a whole, the explanation (which I will treat as a single explanation which is probably clearest in Wang) appears to say that compression is favored in a small chamber, and that rarefaction is favored in a large chamber. This explanation seems to have been the result of a misunderstanding of a simple concept in physics.

According to Boyle's Law (Boyle,1662), in a closed chamber, pressure is inversely related to the volume of gas in that chamber. This seems to have led Wang, Greenberg and Ladefoged to conclude (correctly) that it is easiest (requires the least change in size) to produce a compression in a small chamber, and to conclude (incorrectly) that it is easiest to produce a partial vacuum in a large chamber. The error of this conclusion can be demonstrated easily. If a small plunger is attached to, let us say, the Houston Astrodome and the plunger is pulled out one centimeter, the vacuum produced will be minimal, and insufficient to produce a sound when a door is opened. If the same plunger is now attached to a small pill bottle and the plunger is again pulled out one centimeter, a reasonable popping sound should occur when the bottle is opened. It takes the same effort to produce a rarefaction or a compression in a chamber of the same size. This means that the explanation for ejectives should hold for implosives as well, and that both types of sounds should have the same preferences for place of articulation. This is obviously not the case.

There are two explanations which I would like to suggest, although I do not believe that we presently have sufficient data to choose between them. The first explanation assumes that the only action to change air pressure involves the movement of the

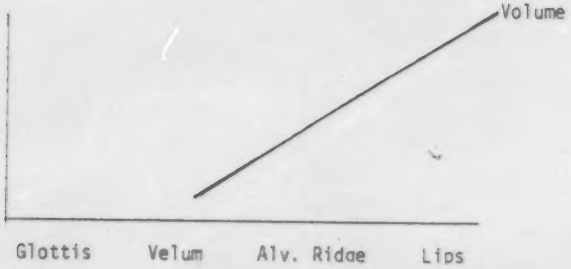
larynx, with some accompanying changes in the diameter of the pharynx. Ejectives, under this view, can still be explained in terms of air pressure, since it is easier to produce a compression at the velum than at the alveolar ridge, and easier to produce a compression at the alveolar ridge than at the lips. In the case of implosives, physiological factors take over. Ladefoged (1968) has noted that the hyoid bone, which is attached to both the larynx and tongue root, moves downward to a considerable degree during implosives, presumably because it is pulled down by the lowering larynx. Such downward movement of the hyoid might be antagonistic to a velar closure, since lowering the hyoid, and simultaneously raising the tongue root, might be extremely difficult given that the two structures are attached. The more limited raising of the tongue body required for alveolars and palatals would be somewhat affected by the lowering of the hyoid, although the effecc would not be as great as that for velars. Finally, labial implosives, not requiring tongue raising at all, would not be adversely affected at all by the actions of the hyoid. This would account for the observed hierarchy, since there would be greater antagonism between the lowering hyoid and the movements of the tongue, as one moved progressively away from a labial articulation.

If this is correct, we should also revise the explanation for ejectives, in that the raising of the tongue body would probably produce a pull on the larynx that would help it rise in the case of velars and alveolars. Since the air pressure preferences work in the same direction, it would seem that the two mechanisms, tongue-pull and pressure, are working in concert to produce the observed preferences. In the case of implosives, the tongue-pull mechanism and the air pressure mechanism work against each other, with the tongue-pull winning out.

This explanation seems plausible, although it may have some problems. If there are two mechanisms working together in ejectives to create the observed pattern, and two mechanisms working against each other for implosives, one would expect the pattern for ejectives to be more pronounced than the one for implosives. This is not the case, as we saw earlier. The pattern for implosives is quite pronounced, while the one for ejectives is not. Furthermore, Ladefoged's data showing that some labial implosives are velarized (Ladefoged,1968) casts some doubt on the explanation. One would not expect velarization, which involves tongue-raising, to occur as a secondary articulation in a sound which made it difficult to raise the tongue body.

Another explanation resembles the Wang-Greenberg-Ladefoged explanation in that it involves the different size cavities created by closures at the different places of articulation, but it adds crucial arguments involving differences in the ability to change the size of those cavities. A plot of the size of the different cavities created by closure at different places in the mouth would look something like the following:

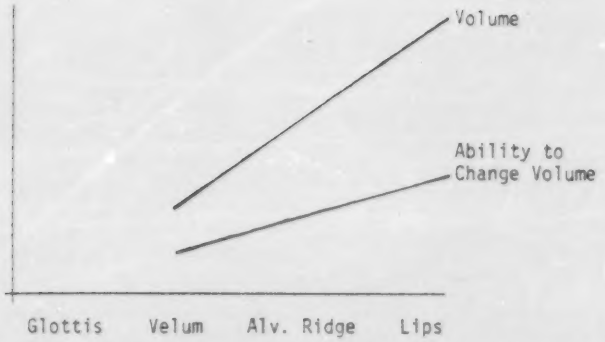
Fig. 1



There are data suggesting that the lowering of the larynx is not the only mechanism available for changing the size of the oral cavity during implosives. Although the data come from experiments with voiced stops, these also involve some expansion of the oral cavity and may allow us to make some inferences. Bell-Berti (1975) noted that the vocal tract expands for voiced stops not only by the lowering of the larynx, but also by the widening of the pharynx and the raising of the velum. Furthermore, Bell-Berti suggests that data from Fujimura et al (1973) which shows a greater area of contact for /t/ than for /d/, may be associated with an expansion of the oral cavity during voiced stops.

It is very likely that implosives, requiring a greater expansion of the oral cavity than voiced stops, use at least the same mechanisms used by voiced stops for increasing the size of the tract. Implosives would be likely to use not only the lowering of the larynx, but also mechanisms for expansion further forward in the mouth. Some evidence for this comes from Greenberg (1970) who notes that implosives show a strong tendency toward apical rather than blade articulation, and also toward retroflexed articulation. This suggests that the parts of the tongue not forming closure are being lowered to expand the oral cavity. In addition, at least in my own production of labial implosives, there is invariably a lowering of the jaw, which would also increase the size of the cavity. However, not all of these mechanisms for increasing the size of the oral cavity are available for all implosives. An implosive produced at the velum would only permit the expansion of the pharynx and the lowering of the larynx, since a raising of the velum would simply force the tongue to move to a higher position. Thus the raising of the velum is only available to places of articulation further front than the velum. Lowering of part of the tongue blade, as a mechanism for increasing the oral cavity, is only available to articulations at the alveolar ridge or further forward. Finally, labial implosives can involve the lowering of the jaw, which will cause a considerable change in the size of the oral cavity. We can plot the ability to change the size of the cavity against the size itself:

Fig.2

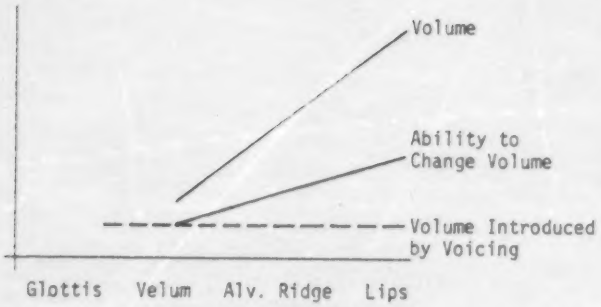


It can now be seen that the change in pressure will be greatest at places of articulation further back in the vocal tract. This is because the ability to change the size of the oral cavity is a larger fraction of the size of the cavity at places relatively far back in the oral tract. Note that this means that the oral cavity has a greater potential to create a pressure change with articulations closer to the glottis. Taking another point of view, for a given change in pressure, the articulatory organs will have to use a smaller percentage of their capacity to change oral cavity size. Presumably, this means that it would be easier to form an ejective relatively far back in the mouth.

This explanation, like the one suggested by Wang and taken up by Greenberg and Ladefoged, might seem to run into trouble with the implosives, since, once again, it would seem that a greater amount of pressure change would be achievable at the back places of articulation. However, for implosives, voicing is almost always present, and this changes the situation.

First, let us note that implosives involve the production of a stop closure with oral pressure relatively close to outside air pressure (Ladefoged,1971). Voicing during an implosive allows air to enter the oral cavity. Since the pressure is close to outside air pressure, the expansion of the oral cavity and the entrance of air into the cavity must cancel each other out. However, the effect of voicing should be approximately the same for all places of articulation. I will plot this effect, together with the size of of the different cavities and the ability to change the size of the different cavities:

Fig. 3



If these relationships hold, velar implosives can barely maintain a relatively low pressure. Alveolar implosives would have a much easier time, and labials would have the easiest time of all. In other words, it may take most of the capacity of the oral cavity for change to maintain pressure equal to outside air pressure during velar implosives. It must take a smaller proportion of the ability for change to do the same for alveolars and labials. This should mean that front articulations are easier to produce than back articulations for implosives.

It is difficult to choose between these two explanations. The second explanation is supported by the counter-examples noted by Campbell (1973), which all involve voiceless uvular implosives. The explanation predicts that any voiceless glottalic consonant, be it implosive or ejective, will favor a back articulation, so that uvular implosives confirm the explanation as long as they are voiceless. The tongue-pull explanation favors ejectives toward the back and implosives toward the front regardless of voicing, so that it should disfavor an uvular implosive. Both explanations suffer from Ladefoged's data showing that the Igbo labial implosive involves velarization. Within the tongue-pull explanation, as already mentioned, one would not expect velarization of a sound that disfavors raising of the tongue body. The second explanation is also disconfirmed by velarization since it assumes that there will be a downward movement to increase the size of the oral cavity rather than the upward movement involved in velarization. These questions will not be answered until the physiological facts of tongue-pull and the aerodynamic facts are better known. Until that time, however, we can at least make sure that any proposed explanations do not ignore the elementary facts of physics.

References

Bell-Berti, F. 1975. Control of pharyngeal cavity size for English voiced and voiceless stops. Journal of the Acoustical Society of America 57.2:456-461.

Boyle, R. 1662. A defence of the doctrine touching the spring and weight of the air. The works of the honorable Robert Boyle, edited by Thomas Birch, London 1744, Vol. 1, p.100.

Campbell, L. 1973. On glottalic consonants. International Journal of American Linguistics 39: 44-46.

Fujimura, O., R.F. Tatsumi and R. Kagaya. 1973. Computational processing of palatographic patterns. Journal of Phonetics 1:47-54.

Greenberg, J.H. 1970. Some generalizations concerning glottalic consonants, especially implosives. International Journal of American Linguistics 36:123-145.

Ladefoged, P. 1968. A phonetic study of West African languages. Cambridge, University Press.

_____. 1971. Preliminaries to linguistic phonetics. Chicago and London. University of Chicago Press.

Wang, W. S-Y. 1968. The basis of speech. Project on Linguistic Analysis Reports, second series, No. 4. Dept. of Linguistics, University of California, Berkeley.

Phonological Notations as Models[1]

John J. Ohala

Abstract

Phonologists often express dissatisfaction with feature systems which do not represent phonological processes or patterns naturally or in self-evident ways (cf. Chomsky & Halle, The Sound Pattern of English, pp. 306-7, 400ff.). This shows that they expect the notation system used to actually model (i.e., incorporate elements isomorphic with) the relevant aspects of the phonological process. As models of sound patterns, currently used notations, i.e., two-dimensional feature matrices, whether or not they are binary or supplemented by markedness conventions, are deficient in some important ways: they have poorly developed formal primitives (the rules of interaction between the substantive primitives) and the substantive primitives are not orthogonal (independent) of each other, or, equivalently, there is no clear separation between primitive entities and derived entities. A simple model which incorporates well-known principles of aerodynamics is presented which satisfies these meta-theoretical requirements and makes useful predictions and explanations about the tendencies of various types of segments to become devoiced, fricated, or aspirated.

Phonological notations need major overhaul. I would like to suggest how this can be successfully done. My argument has four parts.

First, I would like to make a simple distinction -- well known in semiotics -- between what I will call notations as symbols and notations as models. Notations which merely symbolize the thing they stand for are in all respects arbitrary in their form except insofar as they must be differentiated from other symbols used in the same context. Notations which model the thing they stand for are, in at least one respect, non-arbitrary in their form, that is, in some way they are or purport to be isomorphic with the thing they represent. A few simple examples of these two types of notations are given in Figure 1. In the case of the notations under the heading 'symbols,' there is no recognizable part of the symbol that is isomorphic with any part of the entities they stand for. Only in Playboy cartoons and for humorous purposes is any isomorphism suggested to exist between the scientific symbols for 'male' and 'female' and actual males and females. Likewise the graphic symbol '~' does not have seven distinct

---

1. This is a revised version of a paper delivered at the 12th International Congress of Linguists, 29 August--2 September 1977, Vienna, Austria.

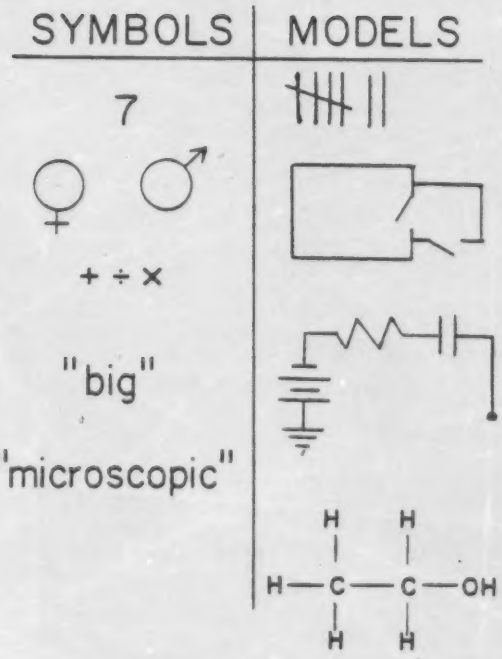


Figure 1.

parts, although the tally mark for '7', on the other side, does. Words, of course, are well known as arbitrary symbols for the things or concepts they stand for. Thus, we do not find it inappropriate that the word 'big' is so small nor that the word 'microscopic' is relatively large. Most mathematical notations are symbols in this sense. However, as for the notations under the heading 'models,' there are obvious respects in which they are isomorphic with the entites they represent. Architects' floor plans are isomorphic with the geometry and proportions of the building they represent. The electrical circuit diagram is isomorphic with the topological arrangement and connectedness of the various components. And so on.

What we require of a symbol, then, is very simple. What we require of a model -- by common practice -- is much more, especially in the case of the more elaborate, more complex models used in modern scientific work. Specifically, we expect a model to represent all and only the facts and relations present in the system it stands for so that some of the behavior of the system will be self-evident, i.e., explained.

'Facts' and 'relations' refer to the usual two parts of a model, substantive primitives and formal primitives, respectively. The substantive primitives or 'ingredients' of the electrical circuit, for example, would be the constituent resistors, capacitors, power sources, etc. The formal primitives or 'rules of operation,' not represented in Figure 1, would be the rules of current flow, and so on, e.g., Ohm's law, which may have a graphical or mathematical expression.

'All and only' means the representation should be minimally redundant. Naturally primitives irrelevant to the particular task should be omitted from the representation. In addition, all primitives that are mentioned should be completely independent of each other, i.e., 'orthogonal' with respect to each other. A simple corollary of this last point is that primitive entities and derived or predicted entities should be kept strictly apart (cf. Vennemann and Ladefoged 1973).

The ultimate purpose of the model, of course, is to explain or make self-evident some behavior of the system by reference to these primitive entities. Ideally the model should also suggest how it may be tested by making predictions about behavior or events that we have not yet observed but could observe given the proper circumstances.

The second point in my argument is to show that today's phonological notations are generally judged by how well they meet the criteria for models.

In an early paper Halle (1962) quite rightly rejected the use of IPA-type notation in the representation of phonological processes because they would not render self-evident the greater naturalness of (1), for example, as opposed to (2). Feature notation of the same processes (3,4) were favored since the greater naturalness of the one process was reflected in the lesser graphical complexity of the representation and by showing (1) to

be assimilatory in character.

(1) a ⟶ ə / __ i, e, ə

(2) a ⟶ ə / __ i, p, z

(3) a ⟶ [-back] / __ [V, -back]

(4) a ⟶ [-back] / __ [V, +high, -back], [C, -sonorant, +anterior, -coronal, -voice, -continuant], etc.

Likewise, the Jakobsonian feature system was criticized because it suggests that the very common change [si] ⟶ [ʃi] is one of dissimilation (cf. 5), whereas it is well-known to be a case of assimilation, as the new Chomsky and Halle (1968) features accurately indicate (see 6).

(5) [C, +continuant, -voice, +diffuse, etc.] ⟶ [-diffuse] / __ [V, -grave, +diffuse].

(6) [C, -high, +anterior, etc.] ⟶ [+high, -anterior] / __ [V, +high, -anterior, etc.]

Some writers have criticized various of the current feature systems because of their non-orthogonality (Liljencrants and Lindblom 1972). The vowel features [low] and [high] of the Chomsky and Halle feature are highly non-orthogonal because a '+' value for one necessarily implies a '-' value for the other. The features [sonorant] and [voice] are non-orthogonal for other, more interesting, reasons (Ohala 1972).

Other examples could be given but I hope this brief list will suffice to show that phonological notations are expected to model, not merely symbolize, phonological processes. None of the above criticisms of feature systems would make sense if the notations were only symbols of phonological phenomena.

The third point is to demonstrate that the potential of current phonological notations to adequately model and thereby explain a vast amount of interesting phonological behavior is very poor. Chomsky and Halle (1968: 400ff.) note this themselves when they admit that there is nothing in a representation such as (7) which reveals it to be much more natural than that

in (8). Unfortunately their solution to this is to posit marking conventions

(7) [-sonorant] ⟶ [-voice]

(8) [-sonorant] ⟶ [+voice]

such as (9). But again, there is nothing in the representation of (9) which suggests in a self-evident way that it is any more natural than (10), its opposite. So we are back where we started: without a representation of the

(9) [u(nmarked value) voice] ⟶ [-voice] / $[\frac{\quad}{\text{-sonorant}}]$

(10) [u(nmarked value) voice] ⟶ [+voice] / $[\frac{\quad}{\text{-sonorant}}]$

process which makes its greater expectedness understandable. The marking conventions are nothing more than an admission that the existing phonological notation is inadequate for the tasks it is used for.

What is the basic fault of current phonological notations, i.e., the various feature systems? Wrong choice of features? There certainly have been some bad choices for features (e.g., 'heightened subglottal pressure') but I don't believe that is the most serious problem facing such systems. In fact, the choice of many of the Jakobsonian features show considerable insight into the way speech sounds are structured (cf. Jonasson 1971). Many of the features proposed are non-orthogonal with respect to each other and this is a shortcoming, but not a very serious one I would think. The most devastating failure of the various distinctive feature systems, in my opinion, is that they have only a very rudimentary set of formal primitives. The list of features themselves, the substantive primitives, are not a bad start for the representation of sound patterns but to constitute a good model they must be supplemented by the rules of interaction of these parameters. It must be possible to show, for example, in a principled way why [-sonorant] segments, i.e., obstruents, tend to be [-voice].

Is there any solution to this problem? Yes, lots of them and this is the final point of my argument. Fant (1960), Flanagan, Ishizaka, and Shipley (1975), Stevens (1971, 1972), Liljencrants and Lindblom (1972), Lindblom (1975) and others have published many interesting and empirically justified models of the speech production system including accounts of the relation between speech articulation and speech acoustics. These models explain or are capable of explaining a great variety of sound patterns (cf. Ohala 1974a,b, 1975a, Ohala and Lorentz 1977).

One quite humble contribution in this area is my own model of aerodynamic events in speech (Ohala 1975b, 1976). Skipping details which can be found in the papers referred to, the model can be characterized as having such substantive primitives (all orthogonal with respect to each other) as: mass of air in the lungs and oral cavity, volumes of the lung cavity and

and oral cavity, area of the glottis and the oral constriction, magnitude of pulmonic expiratory force, and such formal primitives as a version of Boyle's law which determines the air pressure in a cavity as a function of the mass of air involved and the volume of the cavity, an equation which determines air flow through a constriction as a function of the area of the constriction and the pressure drop across it, and so on. From such parameters ('features', if one insists) presence or absence of voicing and noise and their intensities can be predicted (see also Scully 1975, Flanagan, Ishizaka, and Shipley 1975). All of this is incorporated in a computer program.

The model can explain certain universal tendencies in the behavior of speech sounds, e.g., that obstruents tend to become voiceless unless something is done to maintain oral pressure at a lower level than subglottal pressure, that is, to maintain a sufficient transglottal pressure drop. (In this case the model simply derives in a formal way what was well known previously.) It also explains why stops before high close vowels have a greater tendency to be aspirated or affricated than those before non-high vowels. This latter tendency is manifested in Japanese, among other languages. It also predicts that high close vowels or glides, expecially if palatal, may become noisy (fricated) as is now happening in modern Stockholm Swedish.

The model also predicted something which at the time of its making I had not yet observed, namely, that vowels or glides should not become fricated if there is accompanying nasalization. The reason for this is obvious: frication requires a high rate of air flow through a narrow channel which in turn requires a fairly high pressure drop across this channel; the high pressure needed in back of the oral constriction will not occur if the air vents through the nasal cavity. Is this prediction borne out? Yes. In Japanese /h/ is manifested as [ç] before the palatal vowel /i/, e.g., /hikari/ 'light' is [çikaɾi], except when it appears in a heavily nasalized environment, e.g., /hiNhiN/ 'horse's heigh' (onomatopoetic) is [çĩŋhĩŋ]. Likewise, in my dialect of American English /h/ is rendered as [ç] before the palatal glide /j/ except when flanked by nasalized segments, in which case it is [h] (or possibly a voiceless nasal), thus: 'ill-humor' is [ɪlçjumɚ], but 'unhuman' is [ənhjũmn]. Naturally, we should not expect nasalization to interfere with glottal or pharyngeal frication since the air pressure built up behind their constrictions could not vent through the velopharyngeal opening. This is explicable only if the physical structure of the vocal tract is taken into consideration.

The model just discussed can only account for sound patterns which arise due to aerodynamic (and some simple anatomical) factors in speech. For other sound patterns other models are needed, e.g., those referred to above. Of course, a model can only tell us if certain behavior of the system is logically derivable from what we know about the makeup of the system (the primitives). Inevitably there will be some behavior due to as yet unknown factors; no model can or should predict behavior not due to the primitives built into it. In such cases, however, models serve the very useful function of revealing to us exactly how much we don't know about the system.

My message, then, is this: Existing feature systems are very poor at doing what we expect of them: to make the naturalness of certain sound patterns self-evident. The solution to this problem does not lie in the kind of piecemeal revisions that fill the current phonological literature, e.g., adding a feature here, deleting a feature there, using unary or n-ary instead of binary features, relaxing the requirement of simultaneity of features within a segment. A major overhaul is needed. We should start making use of the many already available physical models of the speech system.

Someone may object that what I have advocated as regards representation of phonological phenomena is suitable for purely physical phonetic facts but modern phonological representations reflect what is in the speaker's head, i.e., what he 'knows' about sound patterns. Nonsense. There is no evidence or plausibility for the claim that naturalness and universality of certain sound patterns is due to psychological as opposed to physical factors. Universal sound patterns are universal because physics and human anatomy and physiology are universal.

Acknowledgement

This work was supported in part by the National Science Foundation.

References

Chomsky, N. and Halle, M. 1968. The sound pattern of English. New York: Harper and Row.

Fant, G. 1960. Acoustic theory of speech production. The Hague: Mouton.

Flanagan, J.L., Ishizaka, K., and Shipley, K.L. 1975. Synthesis of speech from a dynamic model of the vocal cords and vocal tract. Bell System Technical Journal 54.485-506.

Halle, M. 1962. Phonology in generative grammar. Word 18.54-72.

Jonasson, J. 1971. Perceptual similarity and articulatory reinterpretation as a source of phonological innovation. Quarterly Progress and Status Reports, Speech Transmission Laboratory, Royal Institute of Technology, Stockholm. STL-QPSR 1/1971.30-41.

Liljencrants, J. and Lindblom, B. 1972. Numerical simulation of vowel quality systems: the role of perceptual contrast. Language 48.839-862.

Lindblom, B. 1975. Experiments in sound structure. Plenary address, 8th International Congress of Phonetic Sciences, Leeds.

Ohala, J. J. 1972. How to represent natural sound patterns. Project on Linguistic Analysis Reports (Berkeley) 16.40-57.

_____ 1974a. Experimental historical phonology. In J. M. Anderson and C. Jones (eds) Historical linguistics II: theory and description in phonology. Proceedings of the 1st International Conference on Historical Linguistics, Edinbrugh. Amsterdam: North Holland. Pp. 353-389.

_____ 1974b. Phonetic explanation in phonology. Papers from the parasession on Natural Phonology. Chicago Linguistic Society. Pp. 251-274.

_____ 1975a. Phonetic explanations for nasal sound patterns. In C.A. Ferguson, L.M. Hyman, and J.J. Ohala (eds.) Nasalfest. Papers from a symposium on nasals and nasalization. Stanford: Language Universals Project. Pp. 289-316.

_____ 1975b. A mathematical model of speech aerodynamics. In G. Fant (ed.) Speech Communication. Proceedings of the Speech Communication Seminar, Stockholm, 1-3 August 1974. Vol. 2: Speech production and synthesis by rule. Stockholm: Almqvist and Wiksell. Pp. 65-72.

_____ 1976. A model of speech aerodynamics. Report of the Phonology Laboratory (Berkeley) 1.93-107.

Ohala, J.J. and Lorentz, J. 1977. The story of [w]: an exercise in the phonetic explanation for sound patterns. Proceedings of the Third Annual Meeting of the Berkeley Linguistics Society 3.577-599.

Scully, C. 1975. A synthesis study of aerodynamic factors in speech segment durations. In G. Fant (ed.) Speech Communication. Proceedings of the Speech Communication Seminar, Stockholm, 1-3 August 1974. Vol. 2: Speech production and synthesis by rule. Stockholm: Almqvist and Wiksell. Pp. 227-234.

Stevens, K.N. 1971. Airflow and turbulence noise for fricative and stop consonants: static considerations. Journal of the Acoustical Society of America 50.1180-1192.

_____ 1972. The quantal nature of speech: evidence from articulatory-acoustic data. In E.E. David and P.B. Denes (eds.) Human communication: a unified view. New York: McGraw-Hill. Pp. 51-66.

Vennemann, T. and Ladefoged, P. 1973. Phonetic features and phonological features. Lingua 32.61-74.

Southern Bantu vs. the world: The case of palatalization of labials

John J. Ohala

Introduction.

Applying typological data to the analysis of particular languages is not new nor is the general principle behind it: the inductive method. Having observed a particular pattern in many languages, positing it in yet another language (given appropriate circumstances) is not wholly unjustified. The purpose of this paper is to demonstrate that one can do a much better job of linguistic analysis if one uses a combination of the inductive and deductive methods, i.e., if one's expectation of a pattern is determined not only by the fact that it has been encountered previously but also because one knows the underlying principles which give rise to it. To illustrate this, I will examine the so-called palatalization of labials, i.e., the shift in place of articulation of palatalized labials or labials followed by a palatal (off-) glide to dentals, alveolars, or palatals (henceforth, for ease of reference, simply 'dentals').

The inductive approach.

A survey of the phonologies of many languages around the world turns up quite a few independent cases of the type of sound pattern exemplified in (1).

(1) [pj, p^j] → [t, ts, tʃ]
[bj, b^j] → [d, dz, dʒ]
[mj, m^j] → [n, ɲ]
etc.

(Occasionally, but not necessarily, intermediate stages may be found, e.g., [ptʃ], [bdʒ], [mn].) Some examples of this pattern are listed in (2) through (10).[1]

(2) Czech (data from Bělič 1966 and Andersen 1973).

Standard Czech	East Bohemian	English gloss
[mjɛstɔ]	[nɛstɔ]	'town'
[pjɛt]	[tɛt]	'five'
[pji:vɔ]	[ti:vɔ]	'beer'
[pjɛknjɛ]	[tɛknjɛ]	'nicely'

1. Paper published in Proceedings of the Fourth Annual Meeting of the Berkeley Linguistics Society, 1978, pp.370-386.

(3) Tai (data from Li 1977).

Siamese	Lungchow	T'ien-chow	English gloss
plaa	pjaa	čaa	'fish'
plau (plaaw)+	pjau	čuu	'empty'
phaai (phaj)+	phjaai	čaai	'to walk'

(+ Accepted current phonemic transcription in Thai.)

Evidently the post-consonantal /l/ changed to /j/ first, then /p (h) j/ changed to the palatal affricate. This same pattern of development is also attested in the Romance languages; see (5), (6) below.

(4) Tibetan (data from Thomas 1948, Benedict 1972, and Chang and Chang 1975).

Old Tibetan	Tzu-ta	Wassu	Mi-li		English gloss
mig ~ myig	temla	temniak	nie +		'eye'
byi-ru	ptsyeru				'coral'
	Gyarong	Lha-sa	Lolopho	Ahi	
bya	pyĕ-pyĕ	ca [tʃa]	byo	do	'bird'
byi-ba		ci-wa [tʃiwa]			'rat'

(+ If Thomas' 'Mi-li' refers to the Tibetan language known as 'Muli', then this word should rather be given as /nöö/ (Nagano 1957).)

(5) Spanish and Portuguese (data from Malkiel 1963).

Latin	Spanish	English gloss
amplu	ancho	'large, spacious'
	Old Spanish	
implēre	(f)enchir	'to fill'
	Portuguese	
plōrāre	chorar	'to weep'
flamma	chama	'flame'
plānu	chão	'floor; level'
plumbu	chumbo	'lead (metal)'

Other data provide evidence for the following separate stages of development of the Latin pl- cluster in these languages: pl- > pj- > tʃ- > ʃ-.

(6) Italian (data from Jaberg and Jud 1928-1940; transcription simplified).

Roman dialect	Genoese and neighboring dialects	English gloss
[pjeno]	[tʃena]	'full'
[pjanta]	[tʃanta]	'to plant'
[er fjato]	[uʃa]	'breath'
[bjanko]	[dʒanku]	'white'

(7) French

Latin	French	English gloss
sapius	sage [saʒ]	'wise'
rubeus	rouge [ʁuʒ]	'red'
rabies	rage [ʁaʒ]	'rabid'
cavea	cage [kaʒ]	'cave'
Proto-Germanic		
laubja	loge [loʒ]*	'arbor'; 'small house'

(* Cf. English 'lobby' and 'lodge', the first having a Germanic origin, the second French.)

(8) Bantu (data from Guthrie 1967-1970).

Proto-Bantu	Tonga	Xhosa, Zulu		English gloss
*pia	phya	-tʃha		'new'
	Kaonde	Sena	S. Sotho	
*biad	-βyal-	-bzaʈ	dʒal	'plant'
	Venda			
*piu	tswhu			'knife'

(9) Classical Greek (data from Meillet and Vendryes 1924).

Pre-Classical Greek	Classical Greek	English gloss
*g^{w}am-yo	βαίνω (cf. Latin venio)	'I come'
*kom-yo- (cf. Latin cum)	κοινός	'common'
*χαλεπ-yω	χαλέπτω	'provoke'
*θαφ-yω	θάπτω	'bury'

(10) Gwari (No. Nigeria) (data from Hyman and Magaji 1970).

Kuta	Ganagana	Nupe	English gloss
byè	dywè	dzò	'sow'
byì	dywí	dzú	'bury'
ọ́pyá	ẹ̀pʃà	ẹ́tswà	'moon'

The deductive approach.

To understand why such changes occur, it is instructive to examine the spectrographic pattern of palatalized labials and compare this with the patterns of plain labials and dentals. Figure 1 shows tracings of spectrograms (the originals published by Fant 1960) of the Russian CV syllables [ba], [bʲa], and [da]. In examining this figure it is necessary to keep in mind the fact that place of articulation cues reside in the second formant (F2) transitions and in the noise bursts. That being the case, it is of interest to see in the figure that the F2 transition for the palatalized labial is more similar to that for the dental than it is to that for the plain labial.[2] Undoubtedly in this instance the noise burst from the release of the stop is a sufficient cue to the labiality of the palatalized labial in spite of the dental-like F2 transition. If a listener were to miss the noise burst cue, however, the consonant would very likely be taken for a dental. Moreover, the impression that such stops were dentals or palatals would be reinforced by any fricative noise generated from the rush of air through the narrow palatal constriction. A sound mistaken for a dental or palatal is likely to be repeated as such. Thus a sound change could occur.

But why should the palatal constriction, a secondary articulation, have a greater influence on the consonantal F2, than the labial constriction, the primary articulation? The beginnings of an answer to this question can be seen in the nomogram in Figure 2 (again, from Fant 1960). The nomogram shows the formant frequencies that would be produced as one varies both the position of the constrictions in the vocal tract and the accompanying lip opening. As can be seen, although the F2 frequency is generally susceptible to change due to both variations in place of constriction and variations in lip opening, its frequency due to a constriction in the palatal region (see arrow) is largely independent of the lip opening. A palatal constriction, even though a secondary articulation, will be the primary determinant of the F2 frequency and will produce a frequency much like that of a dental consonant.

In the case of nasal consonants, there are related but slightly different reasons why the nasal murmur of a palatalized [m] or an [m] coarticulated with the palatal vowel [i], would be acoustically similar to an [n] or [ɲ] (Ohala 1975).

The acoustic similarity of palatalized labials (or labials followed by or coarticulated with palatal vowels) and dentals is

Figure 1. Tracings of spectrographic patterns for the Russian syllables [ba], [bʲa], and [da] (from Fant 1960).

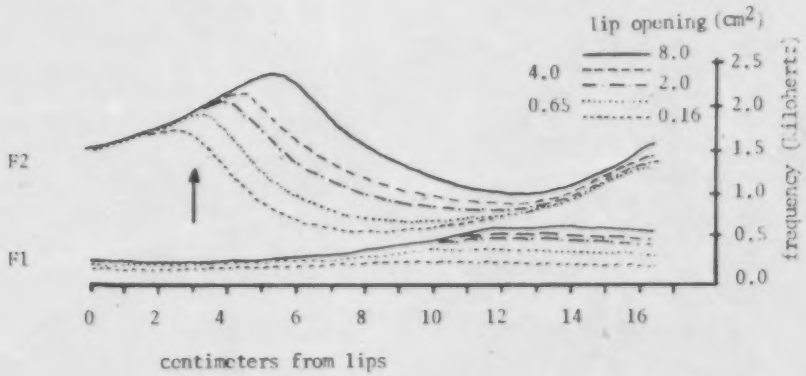


Figure 2. A nomogram (from Fant 1960) showing the frequencies of the first two formants (F1 and F2) that would result from constrictions in various places in the vocal tract (horizontal axis) and from variations in the lip aperture (the parameter). The arrow marks the approximate position of a palatal constriction.

also suggested by the results of various speech perception studies. Lyublinskaya (1966) in studying the confusability of VC transitions (i.e., where there were no bursts to aid identification of the consonant) found palatalized labials were up to 30% more likely to be confused with dentals than were plain labials. Winitz, Scheib, and Reeds (1972) published confusion matrices of CV sequences (where C = p, t, k, and V = i, a, u,) obtained under two conditions: stop burst only and stop burst plus 100 msec of the following vowel. The sequence /pi/ was one of the few stimuli that showed a relatively large decrease in identifiability when 100 msec of vowel was included. In accord with the phonological evidence presented above, most of the confusions of this syllable were with the syllable /ti/. (These two studies, as well as that of Wang and Fillmore 1961, also reveal a strong tendency for labials preceded or followed by /u/ to be misheard as dentals. This effect will be discussed further below.)

In addition, the auditory similarity of the syllables [mi] and [ni] has been revealed in other perceptual studies by House (1957), and Gay (1970).

Analysis of the Southern Bantu Palatalization of labials.

We can now attempt to apply this information to an analysis of some rather unusual cases of palatalization of labials in several Southern Bantu languages, e.g., Zulu, Xhosa, Tswana, Sotho, Venda, Pedi. In these languages the palatalization of labials (and some non-labials) plays an important role in a number of morphological processes, all of which, however, are manifested in about three or four types of (surface) phonological environments. (The data to follow on Southern Bantu are taken from Doke 1926, Meinhof and Warmelo 1932, Cole 1955, Tucker 1929, Jacottet 1927, Guthrie 1967-1970).

One of these types is exemplified in (11). Addition of the causative suffix -ya triggers palatalization. Although we might

(11)	Sotho	verb stem + a	causative	English gloss of stem
		ɛta	etsa	'go'
		kɛna	keɲa	'go in'
	but:	ɬapa	ɬatswa	'wash'
	Tswana	-tlhaleɬa	-tlhaletʃhwa or -tlhaletsha	'become wise'
		-nat'eɬa	-nat'etʃhwa or -nat'etsha	'become pleasant'

wonder a bit as to where the w came from in the derived forms of the last three verb stems (a point that will be addressed later)[3], this seems to be a normal development; it parallels the kind of change we have seen before in many other languages.

Another environment for palatalization is that exemplified in (12) viz., addition of the diminutive suffix -ana. This is a

(12) Diminutive formation of nouns by suffixation of -ana.

Sotho	Noun stem	Diminutive	English gloss of noun stem
	lebese	lebesana	'milk'
but:	moriɬi	morits^{h}wana	'saucer'
Zulu	u:pha:phɛ	u:phaʃa:na	'feather'
	u:ʃu:bu	u:ʃudʒa:na	'meal-water'
	iŋk'a:bi	iŋk'atʃ'a:na	'ox'
Tswana	tl^{h}ap'i	tl^{h}atʃ'wana	'fish'
		(dial. tl^{h}ap'ʃana)	

very unnatural and unexpected environment for this process--given what we have reviewed above. There is no apparent reason why the vowel /a/ or any other phonetic property of the suffix /ana/ should lead to palatalization.

Finally,[4] the most unexpected case of palatalization of labials is that triggered by the addition of the passive suffix -wa to verb stems. Examples are given in (13). Not only should a

(13) Passive formation by suffixation of -wa.

Sotho	Verb stem + a	Passive	English gloss of verb stem
	lesa	leswa	'leave'
	rɛka	rɛkwa	'bury'
but:	bɔɬa	boʃwa	'tie on back'
	tseba	tseʂwa	'know'
	thopa	thotʃwa	'capture'
	ɬopha	ɬotʃwha	'heap up'

following /w/ not cause palatalization of labials, it is the one segment most likely to reinforce the labiality of labials.

A digression on [w].

It might be asked how I can justify that last statement given the evidence mentioned above that in at least 3 perceptual studies labials in the environment of the vowel /u/ were often misheard as dentals. The justification is that /u/, although phonetically close to /w/, differs in the important respect that, unlike /w/, in the environment of a labial it need not have a rapid change in formant frequency. Lyublinskaya presents evidence that one of the cues

for place of articulation of a consonant is direction of F2 transition; presumably a negative (downward) transition for labials and a lack of transition for dentals. Thus the lack of strong F2 transitions between /u/ and a labial could lead to the consonant being taken for dental. Since /w/ necessarily has rapid formant transitions this factor should not apply. There is, moreover, phonological evidence that /w/ should enhance the labiality of labials.[5]

A labial (velar) glide (sometimes < u + V) adjacent to non-labial obstruents generally precipitates a shift to the labial place of articulation (if it leads to any change at all). Examples of this are not uncommon; (14) and (15) provide a few of many cases that could be cited.

(14) Classical Greek (data from Meillet and Vendryes 1924, Meillet 1964).[6]

PIE root	Latin	Greek	English gloss
*yēkʷr̥	iecur	ἧπαρος	'liver'
*ekwōs	equus	ἵππος	'horse'
	Sanskrit		
*gʷiwos	gayaḥ	βίος	'life'

(Cf. also, βαίνω in (9) above.)

(15) Gujarati (data from Turner 1921).

Middle Indic	Gujarati	English gloss
dvāra	bār	'door'
dvē	bɛ̄	'two'
-tvana	-pan	(suffix)

References to further such data can be found in Ohala and Lorentz (1977) and Ohala (forthcoming). These same sources as well as Durand (1956) discuss some of the auditory-acoustic reasons for these sound patterns.

Such data would lead us to expect [w] to reinforce the labiality of labials since it has the power to cause non-labials to become labials.

Previous analyses.

There is a fairly extensive literature on these Southern Bantu languages and I have not had access to most of it. Nevertheless, if there has been any analysis which succeeds in pulling all instances of palatalization together under one rule, it has not found its way into the standard reference works on Bantu. Meinhof and Warmelo (1932) and even Guthrie (1967-1970), for

example, two major contributors to the reconstruction of Proto-Bantu, simply list for these Southern Bantu languages (what amounts to) sound changes of the type *p → t / _ j, w; they apparently saw nothing unusual in the presence of the w in the environment. Meinhof does offer an explanation for the w or labialization that often remains after the palatalization of labials occurs:

> A very peculiar process is that by which sounds to some extent exchange their quality, each giving up some of its own and assuming those of the other. Thus in Sotho ʃya becomes swa. The first sound, ʃ, is a voiceless labial fricative, the second y is a lingual (or more accurately palatal) semi-vowel. The first sound becomes s, i.e. a lingual (strictly an alveolar) voiceless fricative, the second becomes w, i.e. a labial semivowel [16].

However, he seems not to have applied this analysis to cases such as, e.g., Sotho boɸa + wa > boʃwa, where there is no surface y to account for the ʃ.

Doke (1926:139ff) attributed palatalization in the passive to the process of dissimilation since, as it happens, most of these Southern Bantu languages do not have (or permit?) sequences of labial + w. In order to avoid this supposedly forbidden sequence (which would result upon addition of the passive suffix -wa to a stem ending in a labial) speakers, he reasoned, must have shifted the labials to dentals.

In 1970 Talmy Givon and Erhard Voeltz recognized the need to unify all the various instances of palatalization in Southern Bantu under one process triggered by a palatal glide. They found evidence for the 'missing' palatal glides in all the relevant cases of palatalization. They incorporated their views and evidence in various lectures (T. Givon, personal communication).

Stahlke (1976) in arguing for the notion of 'segmental fusion' (essentially that expressed by Meinhof in the above quote[7]), cited Tswana date such as that in (11) and (12) and presented evidence that the causative, passive, and diminutive formations could all be accounted for by one basic rule which involved exchange of features between the labial consonant and a following (sometimes reconstructed) palatal segment.

Herbert (1977) took issue with Stahlke's analysis and rather argued that these cases were morphologically not phonologically conditioned (that is, not phonetically based). His arguments were based primarily on three points: a) that the alternations observed are phonetically unnatural, b) that they applied only in certain morphologically, not phonologically-defined environments, and c) that they show many exceptions and much free variation.

In what follows I propose to provide additional evidence for the Givon-Voeltz-Stahlke analysis. I do not dispute Herbert's claim that these alternations are now activated by specific

morphological environments (e.g., diminutive, passive formations) not by phonetic environments. But this is in no way contradictory to the claim that they all had a common phonetically natural origin.

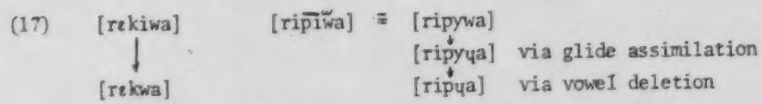
I have already presented evidence for the phonetic naturalness of labials shifting to dentals when followed by palatal glides. What is necessary, then, is to assemble the evidence that palatal glides can be found in earlier stages of the passive and diminutive suffixes.

The history of the passive.

As was pointed out by Stahlke the passive suffix in Tswana has two forms: -wa and -iwa. One can make the case that there was once only one, -iwa. Prosodic factors presumably contributed to the creation of two forms. Zulu, Tswana and Venda for example, use the -iwa form primarily with monosyllabic verb stems but -wa with most polysyllabic stems. We can guess that the -iw coalesced in the latter case, perhaps to [j], perhaps to [ɥ]. Indeed, with simultaneous labialization there is not much difference between these two. As Tucker points out, the labialization of consonants, although indicated in conventional phonetic transcriptions as a w after the sound, is in fact more a prosody of labialization that persists through the segment. If we indicate this prosody as a superscripted line above the consonant, as is done in the Firthian Prosodic tradition, it is easier to see why [j͂ʷ] is equivalent to [ɥ]. Pedi probably still shows some remanents of this early development; see (16).

(16) -rɛka + wa > -rɛkwa 'buy'
-ripa + wa > -ripɥa 'cut'

Assuming that all labial consonants are intrinsically labialized in these languages (which explains why labials cannot support distinctive, i.e., extrinsic, labialization[8]), we may speculate that the two passive forms originated more or less in the sequence indicated in (17).

(17)


(On this point one wonders what use to make of the comment by Jacottet (1927:110) that an older, rarer form of the Sotho passive suffix exists: -uwa, and that it is used mainly after labials? It could validate the claim that labials tend to round the following vowel--especially when it is assisted in this by the -w- --but unfortunately it does not prove that the resulting rounded vowel would be [y].)

Another possible route for the development of a palatal glide after labials is suggested by the passive forms in Venda given in (18) (from Ziervogel and Dau 1961:37).

(18) -bɛba + wa > -bɛbɣwa = [bɛbɣ̈a] ~ -bɛbja 'bear'

The [ɣw] is described as having very little frication.

Since the listeners would expect inherent labialization after the labial b, they may have ignored (i.e., treated as redundant or non-distinctive) the labial part of the following labial velar glide w. What would remain to be considered distinctive in this glide, then, would be the velar component, [ɣ]. This is quite speculative, of course, and, in any case, does not reveal how [ɣ] could change to [j] in the environment given or, indeed, whether the [j] is from the [ɣ].

The history of the diminutive.

In the case of the diminutive formation, as Stahlke pointed out, there is another common form of the suffix which has a palatal initial, -ɲana, and it is possible to trace this and the -ana form to -ɨana. This morpheme exists in many Bantu languages and means 'child'.

Further evidence for the elided j comes from traces of its influence on non-labials, e.g., in Tswana logoŋ 'piece of wood' + dim. suffix > logoɲana. The change of [ŋ] to [ɲ] is plausible only if we assume a palatal glide in the suffix.

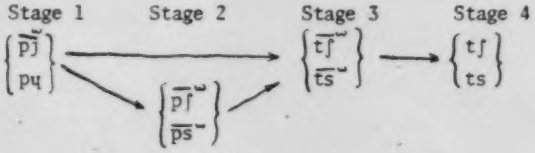
Possible additional evidence for this point comes from the participation of the diminutive suffix in the pattern of vowel harmony in Sotho. In that language it seems that the lower mid vowels shift to higher mid vowels when followed in the next syllable by a high vowel (either i or u) or--and this is the curious part--by labialized consonants formed by passivization or diminutivization, e.g., selepɛ 'axe' + dim. > seletswana (transcription simplified). If we can be sure that simple labialization itself doesn't effect vowel harmony (the evidence isn't very clear on this point, but it seems not to) and if we can trust Tucker's transcription of the Sotho vowels (Cole 1955:xxvii, warns us not to) it would seem that the shift of [ɛ] to [e] in the above example could be attributed to the once-present palatal segment.

The source of the w.

Returning to the question raised earlier, 'where did the w or labialization come from after the palatalization had shifted the labial to a dental?', we can agree with Meinhof and Stahlke that it came from the original labial consonant.[9] From this it follows that the w in the phonetic transcription of passive forms such as tseɣwa has a different source from the w in forms such as rɛkwa. In the former case it comes from the labial in the stem tseba; in the latter

it comes from the passive suffix itself -(i)wa. It is not necessary, though to view the process as entirely one of featural exchange as Meinhof and Stahlke do, at least as far as preservation of the labialization is concerned. It is simpler to assume that labialization spanned the stem-final consonant originally (since, as was mentioned, these labials were probably intrinsically labialized) and it spanned the stem-final consonants after the shift in place of articulation as well. Using the prosodic transcription of labialization, (19) presents a probable scenario for the change, taking up where (17) and (18) left off.

(19)

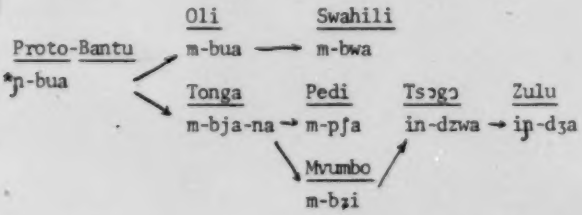


Evidence for Stage 2 comes from, among other things, the passive forms in 'old fashioned' Sotho, e.g., bɔɸa + wa > bɔɸʃwa, as well as some of the variants given above in (12). It should be emphasized that Stage 2 is a possible but not a necessary intermediate stage between Stages 1 and 2. In the perceptual tests reviewed above, subjects' misperceptions, which may be regarded as the stuff sound changes are made from, were abrupt--hearing a p as a t--and did not involve any intermediate stages.

Evidence for the development from Stage 3 to 4, i.e., loss of labialization, was presented earlier (see 11); Cole (1955:43) testifies that among forms showing this variation, those with labialization are older.

Some support for the scheme in (19) comes from the variant reflexes of the Bantu word for 'dog' as given in (20). Here it

(20)



was a palatal nasal noun class prefix ɲ which sometimes exerted a palatalizing influence on the following consonant. The arrows here do not imply direct genetic development but rather that the form near the head of the arrow represents in some sense a further development than the one near the tail.

Conclusion.

Why is it better to combine the inductive and deductive approaches in linguistic analysis? Because they complement each other. In the present case, the inductive approach could tell us that j is a catalyst in the shift of labials to dentals and w helps to shift dentals and velars to labials but it could not tell that in principle it is unlikely that w could also help shift labials to dentals. The deductive approach, in this case reference to the underlying phonetic factors which cause perceptual ambiguity, can tell us which kinds of misperceptions (which might lead to sound changes) are more likely than others. Nevertheless, the deductive approach works perfectly only if our knowledge of the underlying principles of speech production and perception is also perfect. Since this undoubtedly will never be the case, our deductions based on current knowledge may sometimes be erroneous. Only when we find a match between our deductions and inductions can we have some increased confidence that we are on the right track.

Acknowledgements.

The research for this paper was supported by grants from the National Science Foundation and the Committee on Research, University of California. For useful tips and advice I thank Andy Allen, Steve Baron, James Bisso, Talmy Givon, Leon Jacobson, Jeri Jaeger, Ian Maddieson, Yasuhiko Nagano. I especially thank my colleagues Madison Beeler, Carol Justus, and Yakov Malkiel for help with the literature on Greek and Romance. Any errors of fact or interpretation are my responsibility.

Footnotes.

1 In general, I retain the transcription used by the source of the data. This creates potential problems only when palatal glides are sometimes represented as y and sometimes as j. Given the emphasis of this paper and the context of a particular symbol, no ambiguity should arise. When square brackets surround a word IPA transcriptional conventions are followed, i.e., [j] for a palatal glide, [y] for a high front rounded vowel, [ɥ] for a labial palatal glide.

2 The same high F2 is characteristic of labials adjacent to the palatal vowel [i] (Lehiste and Peterson 1961, Fant 1973, Öhman 1966).

3 Leaving aside, for the moment, where the w came from in these forms, the presence or absence of the labialization in the latter two (the Tswana) cases provides a ready explanation for the alternation between the phonetic quality of the resulting affricates [tʃ hw] and [tsh]. As is well known, the presence of lip rounding effectively lowers the resonant frequencies of

the vocal tract. Thus the sibilant noise generated will have a higher or lower center frequency depending on whether or not there is labialization. From such an initial allophonic difference it is quite plausible to find the development of distinctive [ʃ]-type vs. [s]-type fricative releases to the affricates since the primary difference between these fricatives is in their low vs. high center frequencies. In this regard we can note that the lip rounding accompanying English [ʃ] and [tʃ] may not be entirely coincidental: it helps to keep these sounds as distinct as possible from [s].

4 I pass over three other morphological processes that involve palatalization of labials: the formation of the perfect tense of verbs, the formation of the locative of nouns, and the action of the singular prefix of the 5th noun class li > le (and occasionally some other similar prefixes). In general, the analysis of these cases is less controversial than that of the passive and diminutive and, in some cases, was well understood by Bantuists early on (cf. Tucker 1929:85ff).

5 [m] in the environment of [w] or [u] is liable to shift to [ŋ] but not [n] or [ɲ]. The reasons for this, which are relevant only to nasals, are given by Ohala and Lorentz 1977a, b. In view of this it is interesting to note the varying fate of stem-final [m] when it is subject to the same derivations that palatalize the obstruents. The result seems to hinge in part on whether labialization is retained or not. If it is not, we find only [ɲ]; if it is retained we can find either [ɲ] or [ŋ]. Thus, e.g., Zulu int'a:mɔ 'neck' + dim. > int'aɲa:na, but Sotho lelene 'tongue' + dim. > lelenwana. One could venture the prediction that [ɲw] sequences are unstable and will shift either to [ɲ] or [ŋw].

6 Before front vowels, however, PIE labial velars generaly become Greek dentals, e.g., cf. Latin que 'and', but Greek τε. See Allen (1957) for an interesting discussion of this exception and of exceptions to the exception.

7 See also Henderson 1975 for similar views.

8 The Southern Bantu languages are in no way unusual in not having distinctively labialized labials or labials followed by w. Even in English, clusters of labial + w exist only by virtue of some rather uncommon loanwords, e.g., bwana. Many other examples could be cited (see Ohala, forthcoming).

9 Muddying the issue somewhat is the fact that some non-labial consonants also show labialization in these derivations. Such cases, however, can be attributed to a stem-final rounded vowel, e.g., Tswana lekoto 'leg' + dim. > lekotwana.

References.

Allen, W. S. 1957. Some problems of palatalization in Greek. Lingua 7.113-133.
Andersen, H. 1973. Abductive and deductive change. Lg. 49.765-793.
Bělič, J. 1966. K otázce zaniklého litomyšlského t, d, n za někdejší p', b', m'. Slavica Pragensia 8, Acta Univ. Carolinae, Philologica, 1-3.
Benedict, P. K. 1972. Sino-Tibetan: a conspectus. Cambridge Univ. Press.
Chang, B. S. and Chang, K. 1975. Gyarong historical phonology. Bull. of Inst. of History and Philology. Academia Sinica. 46.3.391-524.
Cole, D. T. 1955. An introduction to Tswana grammar. London: Longmans, Green.
Doke, C. M. 1926. The phonetics of the Zulu language. Bantu Studies 2, Special Number.
Durand, M. 1956. Du rôle de l'auditeur dans la formation des sons du language. J. de Psych. Norm. et Path. 52.347-355.
Fant, G. 1960. Acoustic theory of speech production. The Hague: Mouton.
Fant, G. 1973. Stops in CV-syllables. In G. Fant. Speech sounds and features. Cambridge: MIT Press. Pp. 110-139.
Gay, T. 1970. Effects of filtering and vowel environment on consonant perception. JASA 48.993-998.
Guthrie, M. 1967-1970. Comparative Bantu. 4 vols. Gregg.
Henderson, E. J. A. 1975. Feature-shuffling in Southeast Asian languages. Paper presented at 8th Sino-Tibetan Conference, Berkeley.
Herbert, R. K. 1977. Morphophonological palatalization in Southern Bantu: a reply to segmental fusion. Studies in African Linguistics 8.143-171.
House, A. S. 1957. Analog studies of nasal consonants. JSHD 22. 190-204.
Hyman, L. M. and Magaji, D. L. 1970. Essentials of Gwari grammar. Occasional Papers No. 27, Institute of African Studies, Univ. of Ibadan.
Jaberg, K. and Jud, J. 1928-1940. Sprach und Sachatlas Italiens und der Südschweiz. Zurich.
Jacottet, E. 1927. A grammar of the Sesuto language. Bantu Studies 3, Special Number.
Lehiste, I. and Peterson, G. E. 1961. Transitions, glides, and diphthongs. JASA 33.268-277.
Li, F. K. 1977. A handbook of comparative Tai. Univ. Press of Hawaii.
Lyublinskaya, V. V. 1966. Recognition of articulation cues in stop consonants in transition from vowel to consonant. Soviet Physics--Acoustics 12.185-192.
Malkiel, Y. 1963. The interlocking of narrow sound change, broad phonological pattern, level of transmission, areal configuration, sound symbolism. Archivum Linguisticum 15.2.144-173; 3.1-33.

Meillet, A. 1964. Introduction à l'étude comparative des langues indo-européennes. University of Alabama Press.
Meillet, A. and Vendryes, J. 1924. Traire de grammaire comparée des langues classiques. Paris: Librairie Ancienne Edouard Champion.
Meinhof, C. and Warmelo, N. J. V. 1932. Introduction to the phonology of the Bantu languages. Berlin: Dietrich Reimer/ Ernst Vohsen.
Nagano, Y. 1975. Phonemics of the Muli language. Annual Report of Japan Assoc. for Tibetan Studies No 21. Pp. 10-11.
Ohala, J. J. 1975. Phonetic explanations for nasal sound patterns. In: C.A. Ferguson, L. M. Hyman, and J. J. Ohala (eds.) Nasálfest. Papers from a Symposium on Nasals and Nasalization. Stanford: Language Universals Project. Pp. 289-316.
Ohala, J. J. Forthcoming. The origins of sound patterns in language.
Ohala, J. J. and Lorentz, J. 1977a. The story of [w]: an exercise in the phonetic explanation for sound patterns. BLS 3. 577-599.
Ohala, J. J. and Lorentz, J. 1977b. Acoustic reasons why labio-velars are both labials and velars. JASA 61. 932.
Öhman, S. E. G. 1966. Coarticulation in VCV utterances: spectrographic measurements. JASA 39. 151-168.
Stahlke, H. F. W. 1976. Segment sequences and sequential fusion. Studies in African Linguistics 7. 41-63.
Thomas, F. W. 1948. Nam. An ancient language of the Sino-Tibetan borderland. Publications of the Philological Society [of London]. 14.
Tucker, A. N. 1929. The comarative phonetics of the Suto-Chuana group of Bantu languages. London: Longmans, Green.
Turner, R. L. 1921. Gujarati phonology. JRAS, 1921, part 3. Pp. 329-365; 505-544.
Wang, W. S-Y. and Fillmore, C. J. 1961. Intrinsic cues and consonant perception. JSHR 4. 130-136.
Winitz, H. Scheib, M. E., and Reeds, J. A. 1972. Identification of stops and vowels for the burst portion of /p,t,k/ isolated from conversational speech. JASA 51.1309-1317.
Ziervogel, D. and Dau, R. S. 1961. A handbook of the Venda language. Pretoria: Univ. of South Africa.

Explaining the Intrinsic Pitch of Vowels[1]

John J. Ohala

and

Brian W. Eukel

Abstract

The reason why, other things being equal, the pitch of vowels correlates with vowel "height" has long been a subject of speculation. The effect is usually attributed either to acoustic coupling between vocal tract and vocal cords, or to the pull of the tongue on the vocal cords. To determine which of these hypothesized factors is responsible we observed the effect on the pitch of vowels when one of these factors (impedance of the vocal tract) was held constant, while the other (tongue stretch) was varied over an exaggerated range, viz., by having eight speakers utter special test utterances when their jaw was propped open by small wooden blocks ≤10mm thick. (Lindblom and Sundberg have shown that under such conditions speakers can achieve acoustically normal vowels by increasing the distortion -- and presumably the stretch -- of the tongue body with respect to the jaw and larynx-hyoid complex.) We found a small but consistent increase in the pitch interval between high and low vowels when the subjects spoke with propped-open jaw, thus supporting the "tongue-pull" hypothesis.

The systematic variation of the mean fundamental frequency - or pitch - of vowels has been noted for several decades, namely that the pitch varies directly with the "height" of the vowel (Crandall 1925, Lea 1972).

Various hypotheses have been proposed to account for this effect and of these two have survived and are given serious consideration today: We'll call these the "coupling" hypothesis and the tongue-pull hypothesis (cf. other reviews by Atkinson 1973b, Ohala 1972b, Delos et al., 1976).

The coupling hypothesis, advocated by Atkinson, (1973a, 1973b), and Lieberman (1970), suggests that when the 1st formant of the vowel is near the fundamental frequency -- as is the case with higher vowels -- acoustic coupling occurs between the vocal tract and the vocal cords, similar to what happens between the resonances of the bugle and the bugler's lips, namely, the tract resonances effectively "dictate" to the source the frequencies at

---

1. Paper given at 92nd meeting of the Acoustical Society of America, 15-19 November, 1976, San Diego, CA; abstract also in: JASA

which it can vibrate (Benade 1973, Ishizaka and Flanagan 1972).

The tongue-pull hypothesis, offered by Lehiste (1970), and by Ladefoged (1964a,b), among others (although I should add, recently repudiated by Ladefoged; cf. Ladefoged, De Clerk, Lindau and Papçun 1972), suggests that somehow or other, in producing high vowels, the tongue pulls on the larynx and thus affects the tension of the vocal cords and thus the pitch of the voice.

We report here an attempt to test these two hypotheses. This test was suggested to us by Björn Lindblom (personal communication). Simply stated, we sought a circumstance where one of the hypothesized causal factors, vocal tract impedance, was held constant, while the other hypothesized causal factor, degree of tongue pull on the vocal cords, was varied. If the tongue-pull hypothesis is correct, such a circumstance should show systematic variation in the pitch interval between high and low vowels; if the coupling hypothesis is correct, there should be no change in the pitch interval.

Earlier work by Lindblom & Sundberg (1971) showed that speakers can produce acoustically normal vowels when speaking with the jaw in abnormal fixed positions by producing compensatory tongue gestures, e.g., they can produce good [i] and [u] vowels with a propped-open jaw by increasing the elevation of the tongue. The vocal tract impedance, if it varies at all, would be expected to decrease in such a case, not increase, and would thus lead to a lessened interval between the pitches of high and low vowels.

We therefore had 8 adult speakers, 4 males and 4 females, speak English nonsense words of the sort [əmimə], [əmaməl], etc., containing the six stressed monophthongal vowels [i, u, ɛ, ʌ, æ, a], about 16 tokens each, under four conditions: normal jaw mobility, midjaw propped open with small wooden spacers 1mm, 5mm, and 10mm thick, placed bilaterally between the molars. The order of the conditions was varied for each subject so that order effects would cancel out. Speakers were given frequent 5 or 10 minute breaks during the readings. Three speakers read the randomized sequence of words from a printed list, but because the subjects had difficulty figuring out the intended phonetic quality of the vowel from the spelling, the remaining 5 subjects spoke the words as they heard them from a cue tape.

The pitch of all tokens was extracted by a custom-made pitch extractor and written out via an oscillograph. The peak pitch of each token was measured by hand. Each speaker's pitch data was converted to the standard $\bar{Z}$-value. (That is, all data was normalized by replacing values in Hz by their corresponding standard deviation value, the standard deviation computed from all utterances of the given individual, over all four conditions. E.g., if a speaker's peak pitch values had a mean of 100 Hz with a SD of 5 Hz, then a measurement of 95 Hz at 105 Hz would be converted to the normalized values of -1 and +1 respectively.)

The data of one female subject had to be excluded from the study since half way through the reading her pitch level dropped about 10 or 15 Hz and

thus did not lend itself to the simple normalization process we used. The pattern of the male and female speakers were similar enough to justify pooling their data for the final analysis.

Auditory analysis of all tokens and an acoustic analysis of a small random sample of the vowels showed that there was no measureable spectral difference between the vowels spoken under the various conditions.

The results for the 7 subjects are shown in Figure 1. In the figure are shown average normalized pitch for the 6 vowels for the 4 conditions: normal (jaw free to move), and then with jaw spacers 1, 5 and 10mm in thickness.[2] Each data point represents the average of some 100 measurements. In accord with most previous studies, high vowels had higher pitch than low vowels.

The difference in the pitch interval between [u] and [a] in the normal condition and [u] and [æ] in the 10mm condition is significant: $p < 0.05$. Table 1 presents details of the statistics of this data.

We take this as support for the tongue-pull hypothesis and lack of support for the coupling hypothesis.

The lack of difference between this pitch interval in the 1mm condition and that in the normal condition suggests that it is not simply the artificial experimental situation of speaking with something clenched between the teeth that gives rise to the effect. The lack of difference between the intervals in the 5mm condition and the 1mm or normal condition suggests that this degree of stretch did not substantially deviate from normal speech.

There are additional reasons for doubting the coupling hypothesis. First of all, current models of vocal tract--vocal cords interaction suggest that the degree of acoustic coupling is relatively low and in any case would produce an effect such as the intrinsic pitch of vowels only when the vowel Fo and F_1 almost exactly coincide (Ishizaka and Flanagan 1972), which is hardly ever true for most adult speakers. In fact it seems that if the vowel Fo is close to but below F_1, the normal pitch will be lowered. Indeed, Delos et al. (1976), modeling with a computer the tract-source interaction, found a slight inverse correlation between vowel height and vowel pitch.

Moreover, the coupling hypothesis would predict that speech produced in a helium-air mixture, where the formants are raised without affecting the Fo very much, would show a lessened difference between the pitch of high and low vowels. However, Beil (1962) presented data which showed this pitch

---

2. The data for the 'mid' vowels [ɛ] and [ʌ] should be regarded cautiously since speakers' pronunciation of these vowels varied. [ɛ] in particular, ranged in pronunciation from [ɪ] to [æ]. It is generally recognized that [ɪ] and [ɛ] are merging in nasal environments in the speech of many Americans.

Table 1. Each cell gives the mean normalized pitch, the standard deviation (in parentheses), and the population size for the indicated vowel in the indicated condition. Values drawn from seven subjects. Below are some details of the statistics used to figure the significance of the difference in the maximum pitch intervals between the normal condition and the 10mm block condition.

Condition \ Vowel	a	ɛ	ʌ	i	u
Normal	-0.81 (0.77) 99	-0.58 (0.87) 115	-0.65 (0.72) 111	-0.53 (0.99) 118	0.23 (0.86) 114	0.44 (0.82) 116
1 mm	-0.18 (0.74) 109	-0.33 (0.82) 103	-0.15 (0.81) 110	-0.21 (0.78) 109	0.72 (0.82) 106	0.93 (0.77) 113
5 mm	-0.11 (0.87) 115	-0.25 (0.81) 110	-0.17 (0.76) 108	-0.14 (0.71) 115	0.78 (0.89) 103	0.93 (0.91) 110
10 mm	-0.47 (0.79) 111	-0.47 (0.79) 119	-0.29 (0.90) 103	-0.24 (0.86) 107	0.62 (0.97) 114	1.09 (1.00) 108

Maximum Intervals

Normal Condition: $u\text{-}a = Mean_u - Mean_a = 1.25$

$Var\ (u\text{-}a) = Var_u + Var_a = 1.27^*$

$N(u\text{-}a) = 99^{**}$

10 mm Condition: $u\text{-}a = 1.56$

$Var\ (u\text{-}a) = 1.62$

$N\ (u\text{-}a) = 108$

Applying 1-tailed t-test to check significance of the difference between these two maximum pitch intervals:

$t = 1.84$

$df = 2.05$

$p < 0.05$

* Independent variances are additive. ** N is taken as the lesser of the N's of two populations.

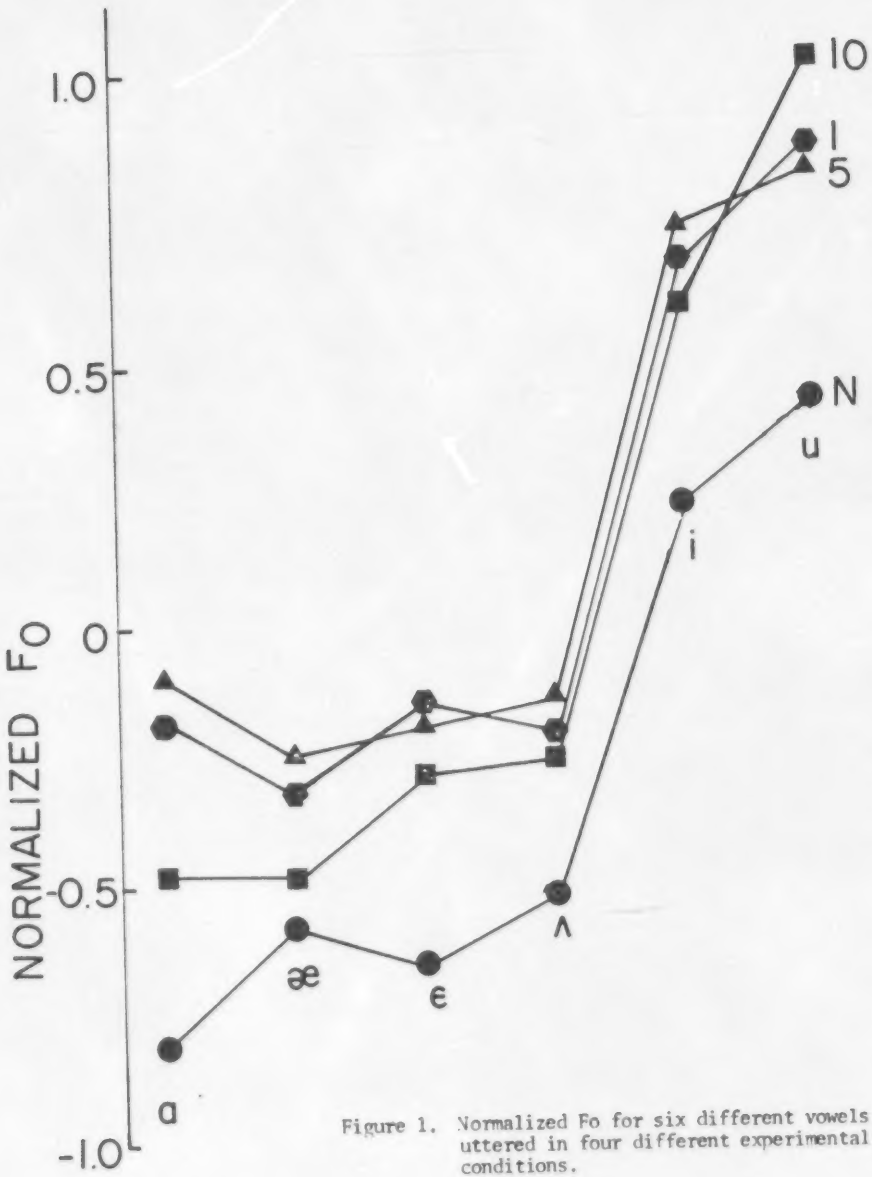


Figure 1. Normalized Fo for six different vowels uttered in four different experimental conditions.

interval is substantially the same in helium and normal speech.

There are also additional reasons for favoring the tongue-pull hypothesis, namely, that although it has been shown that there is generally an inverse correlation between vowel height and hyoid or larynx position (which is why Ladefoged abandoned his support of this hypothesis), laminographic data exists which shows that one feature of the soft tissues of the larynx, ventricle size, does vary directly with vowel height (Shimizu 1960, 1961, van den Berg 1955). [3] We would guess increased ventricle size is a reflection of some factor which could lead to increased pitch. There is, moreover, widespread evidence from x-rays of vowels that shows that the pharynx is wider (tongue root more advanced) in the high vowels [i] and [u], in comparison with the lower vowels (Eijkman 1914, Holbrook and Carmody 1937, Fant 1960, Perkell 1969). Houde (1968) presents evidence that the high position of the tongue during [i] and [u], viz-a-viz [a], actually involves greater physical stretch of the surface tissue on the back of the tongue (as determined by changes in the distance between metal pellets attached to the tongue surface). Plausibly, this stretch could lead to changes in the tension of the soft tissues of the larynx, including the vocal cords themselves. Indeed, there is suggestive evidence that the angle of the epiglottis, and the "internal aspect of the larynx", i.e., how the vocal cords look when viewed from above, also show systematic variation with vowel quality in ways that would be consistent with higher tension for higher vowels (Eijkman 1929, 1933). It is reasonable to suggest that the intrinsic pitch of vowels is a manifestation of the variation of pitch by means of changes in the vertical, not the anterior-posterior tension, which is a mechanism whose existence one of the present authors gave evidence for at a previous meeting of the Acoustical Society (Ohala 1972a).

---

3. Shimizu's data is particularly noteworthy. He showed that the average order for larynx height for the production of the 5 Japanese vowels was, from highest to lowest: [i, e, a, o, u]. This pattern agrees with that found by many other researchers (see Ohala 1972b for a review of this literature) and shows an inverse correlation between vowel height and larynx height, at least for [a, o, u]. It was this inverse correlation which prompted Ladefoged to abandon the 'tongue-pull' hypothesis. Nevertheless, as mentioned, these same 10 subjects showed the average order of ventricle size, from largest to smallest, to be: [u, i, e, o, a], which is directly correlated with the tongue "height" of the vowel.

References

Atkinson, J.E. 1973a. Intrinsic Fo in vowels: physiological correlates. JASA 53.346.

_____ 1973b. Aspects of intonation in speech: implications from an experimental study of fundamental frequency. University of Connecticut, Doctoral Dissertation.

Bernade, A. 1973. The physics of brasses. Scientific American. July, pp. 24-35.

Beil, R.G. 1962. Frequency analysis of vowels produced in a helium-rich atmosphere. JASA 34:347-49.

Berg, J. van den. 1955. On the role of the laryngeal ventricle in voice production. Folia Phoniatrica 7:57-69.

Crandall, J.B. 1925. The sounds of speech. BSTJ 4:586-626.

Debos, M., Guerin, B., Mrayati, M., and Carré, R. 1976. Study of intrinsic pitch of vowels. JASA 29:S72.

Eijkman, L.P.H. 1914. The tongue-position in the pronunciation of some vowels as set forth by röntgen-photographs. Vox, Heft 3, 24 Jahrgang, June, pp. 129-243.

_____ 1929. The area of the glottis in vowels. English Studies 9.2: 49-58.

_____ 1933. The internal aspect of the larynx in speech. ANPE 8-9: 213-219.

Fant, G. 1960. Acoustic theory of speech production. Mouton.

Holbrook, R.T., and Carmody, F.J. 1937. X-ray studies of speech articulation. University of California Publications on Modern Philology 20.4.

Houde, R.A. 1968. A study of tongue body motion during selected speech sounds. SCRL Monograph No. 2.

Ishizaka, K. and Flanagan, J.L. 1972. Synthesis of voiced sounds from a two-mass model of the vocal cords. BSTJ 51:1233-1268.

Ladefoged, P. 1964a. Some possibilities in speech synthesis. L&S, 7:205-14.

_____ 1964b. A phonetic study of West African languages; an auditory-instrumental survey. West African Language Monograph Series, No. 1. Cambridge University Press.

Ladefoged, P., De Clerk, J., Lindau, M., and Papcun, G. 1972. An Auditory-motor theory of speech production. UCLA-WPP 22:48-75.

Lea, W.A. 1972. Intonational cues to the constituent structure and phonemics of spoken English. Purdue University, Doctoral Dissertation.

Lehiste, I. 1970. Suprasegmentals. MIT Press.

Lieberman, P. 1970. A study of prosodic features. HL-SRSR-23-1970:179-208.

Lindblom, B., and Sundberg, J. 1971. Neurophysiological representation of speech sounds. Proceedings of the 15th World Congress of Logopedics and Phoniatrics, Buenos Aires.

Ohala, J.J. 1972a. How is pitch lowered? JASA 52:124.

_____ 1972b. Explanations for the intrinsic pitch of vowels. Unpub. MS.

Perkell, J.S. 1969. Physiology of speech production; Results and implications of a quantitative cineradiographic study. MIT Press.

Shimizu, K. 1960. On the motions of the vocal cords in phonation studied by means of the high voltage radiograph movies. (In Japanese: English summary.) The Oto-Rhino-Laryngology Clinic (Zibi-Inko-ka Rinsyo) 53.4: 446-461.

_____ 1961. Experimental studies on movements of the vocal cords during phonation by high voltage radiograph motion pictures. Studia Phonologica 1:111-116.

Listeners' Ability to Identify Languages by Their Prosody [1]

John J. Ohala

and

Judy B. Gilbert[2]

Abstract

In an effort to see whether there are recognizable language-specific prosodic characteristics, 41 listeners were asked to identify the languages used in each of 54 recorded passages of speech which had been electronically processed to preserve only the fundamental frequency, amplitude, and timing of the original speech (which was obtained from 3 speakers each of the languages English, Japanese, and Cantonese). The overall rate of correct responses was 56.4%, far above the 33.3% rate for pure guessing, indicating that in general subjects were successful at the task. Possible implications and further applications of the test in second language teaching are discussed.

Introduction

It is a common observation that different languages "sound" different, not only due to their use of different segments but also due to differences in their intonation or prosody. If true, this could have considerable importance in many practical areas, including second language teaching,[3] speech synthesis, and automatic speech recognition. Unfortunately most claims of this sort -- whatever their intrinsic merit -- cannot be directly applied to these areas since the claims are usually based on subjective experience and are couched in very impressionistic terms. The purpose of the research to be reported here was to attempt to verify that languages do differ in their prosody by seeing if listeners are successful in identifying the language used by a speaker when they hear only the fundamental frequency, amplitude, and certain timing characteristics of the original voice signal.

Some progress has already been made on this question. Previous investigations on this topic have used a variety of techniques and languages. To

---

1. To appear in: Festschrift for Georges Faure.
2. Department of Linguistics, University of California, Davis, CA 95616
3. We are pleased to dedicate this paper to Professor Georges Faure who has long emphasized the necessity of including prosodic training in second language teaching.

produce a signal in which segmental information has been removed, Atkinson (1968) and Bonte (1975) generated a pulse train having the same frequency and amplitude as the original speech signal, Bush (1967) and Richardson (1973) used low-pass filtered speech,[4] and Maidment (1976) used a laryngographic signal, i.e., an indication of short-term variations in glottal electrical resistance -- a signal, therefore, closely related to the original glottal waveform and uninfluenced by the resonances of the vocal tract or by supra-glottal noise sources. The languages which listeners had to identify in these studies were: English and Spanish (Atkinson), American English, British English, and Indian English (Bush), Black English and White English (Richardson), English, French and Chinese (Bonte), and English and French (Maidment). All of these investigators used predominantly read speech. Although all investigators found better-than-chance levels of language identification, all seem to have shown that it was a difficult task since listeners' average scores could be as low as (approximately) 65% in a task involving the identification of 2 languages (Maidment) or 41% in a task involving 3 languages (Bonte). Other effects emerging from these studies are that longer stretches of speech are easier to identify than shorter stretches and that listeners correctly identify their own language better than other languages.

In the present study we sought to verify that language identification via prosody could be successfully done using conversational speech (i.e., not based on a written text) and to verify in addition, the following hypotheses:

1) Long passages will be better identified than short passages.
2) The listener will be more successful at making the distinction own language/not own language as opposed to the distinction other language$_1$/other language$_2$.
3) Identifications will be improved if listeners are given a prior training session incorporating both the original unprocessed voice signal along with the processed version.

Due to the choice of listeners it turned out that we could also test the additional hypothesis:

4) Bilingual or trilingual speakers (of the languages used in the test) will outperform monolinguals.

For the languages used in the test we chose English, Japanese, and Cantonese, partly because of availability of subjects, but primarily because according to traditional prosodic descriptions they represent three distinct prosodic types: languages using stress, accent, and tone, respectively. In addition, English and Japanese are recognized as having different timing patterns: so-called "stress-timing" and "syllable- (or mora-) timing",

---

4. Cohen and Starkweather (1959) approached the same question but it may be doubted whether their procedure of low-pass filtering speech samples at 600 Hz was sufficient to remove language-specific segmental cues.

respectively. Thus, these languages should be easily differentiated.

Method

Recordings of unrehearsed conversational speech were obtained from 12 adult males, 4 each native speakers of American English, Japanese, and Cantonese. The stimulus material for the test was drawn from the speech of 9 of these speakers (3 of each of the 3 languages); the speech recorded from the remaining 3 speakers was used for a training sequence (see below). The recordings were obtained while each speaker, alone in a sound-treated recording room, responded to comments or questions posed him by another native speaker outside the room (whose voice the speaker could hear over an earphone). Both could see each other through a window in the recording room. From the speech recorded from each of the 9 speakers, 6 fluent passages were selected, 3 "short" passages ($\leq$ 10 sec) and 3 "long" passages ($\geq$ 15 sec, but $\leq$ 25 sec), yielding a total of 9X6= 54 passages (actually, through an error the samples from one Japanese speaker included 4 short passages and only 2 long ones; thus there were a total of 28 short passages and 26 long ones). The fundamental frequency and amplitude of these passages was extracted using a custom-made circuit (Krones 1968) and fed to separate inputs of a voltage-controlled-frequency and voltage-controlled-amplitude signal generator (Wavetek Model 146) which was set to produce a triangular pulse train. When there was no fundamental frequency signal, i.e., when there was no voicing, the amplitude of the processed signal went to zero. In this way the original speech signal was converted to a "buzz" having the same frequency, amplitude, and timing (at least with respect to the relative timing of voice on/voice off).

In addition, a training passage was constructed which included speech samples from the 3 extra speakers, both in their original unprocessed form and in the processed "buzz" form. The training passage plus instructions followed by the 54 test samples in "buzz" form, randomized, were dubbed onto a master stimulus tape for presentation to listeners. Each test item was followed by a short tone marking the end of the item plus 5 seconds of silence during which the language was to be identified. The instructions, which were presented in writing as well as on the tape, were as follows:

> This is a test to find out if people can recognize a language based on the intonation pattern alone. That is, if everything is removed but the "music" of the speech, can the language be identified? An example of intonation patterns in English is the following: 1. 'You're not going?' 2. 'You're not going!' Nothing was changed in those sentences but the intonation pattern.
>
> This test consists of recorded excerpts taken from conversations in Japanese, Cantonese, and English. These excerpts have been electronically processed so that only the musical part, the intonation, remains. The "music" representation sounds like a buzz. We will ask you to listen to this buzz and guess what language it is from. Half of the items are short and half are long

[however, see above]. Each item will be followed by a marker sound, like this [tone] and then you will have a 5 second pause to mark the answer sheet. If you think the buzz was from English, please write 'E' after the item number. If you think the buzz was from Japanese, write 'J'; Cantonese, write 'C'. There are many speakers of the same language included in the samples. The items are in random order.

To help you hear the buzz tone better, we are going to give you a brief training period. You are going to hear a phrase taken from a sentence in English. It will be followed by a buzz representing that phrase. [Samples played.]

Now you will hear the entire sentence, including the phrase, followed by the buzz representation. [Samples played.]

Now you will hear a sample of Cantonese, followed by the buzz representation. [Samples played.]

Now you will hear a sample of Japanese, followed by the buzz representation. [Samples played.]

We want to be sure that you understand the procedure for this test, so we are going to give you some test samples. Please mark your answer sheet 'C' if you think the buzz was Cantonese; 'J' if you think it was Japanese; or 'E' if you think it was English. [Three short samples played.]

You are now ready to begin the test. None of the voices used for this test were included in the examples you heard before.

The stimulus tape was presented individually to 41 adult listeners over a high-quality playback system via earphones. Of the 41 Listeners, 18 were native speakers of American English, 12 of Cantonese, and 11 of Japanese. It was not feasible to be completely selective with regard to listeners' mastery of or familiarity with the other 2 languages used in the test aside from their native language. As it turned out, there were 12 monolinguals, all native English speakers, 25 "bilinguals" (including those who had studied one of the other languages extensively in school), and 4 trilinguals, all among the native Cantonese speakers. The training passage was deliberately withheld from 5 English speaking listeners (4 monolinguals and 1 bilingual); all other listeners heard the training passage. The answer sheet consisted of 56 numbered squares of which only 54 were used.

Results

The results are given in Table 1 through 12. 'E', 'J', and 'C' stand for 'English', 'Japanese', and 'Cantonese', respectively. Listed on the

Table 1. Monolinguals (N=12)

Presented \ Heard	E	J	C
E	140	51	25
J	52	85	79
C	46	69	101

correct: 50.3%

Table 2. Bilinguals (N=25)

	E	J	C
E	280	113	57
J	87	238	125
C	85	76	289

correct: 59.3%

Table 3. Trilinguals (N=4)

	E	J	C
E	34	13	25
J	22	40	10
C	23	8	41

correct: 53.2%

Table 4. English (with training passage) (N=13)

	E	J	C
E	168	39	27
J	50	102	82
C	53	64	117

correct: 55.1%

Table 5. English (no training) (N=5)

	E	J	C
E	49	29	12
J	26	31	33
C	19	32	39

correct: 44.1%

Table 6. Japanese (N=11)

	E	J	C
E	107	74	17
J	39	95	64
C	43	31	124

correct: 54.9%

Table 7. Cantonese (N=12)

	E	J	C
E	130	35	51
J	46	135	35
C	39	26	151

correct: 64.2%

Table 8. Short passages

	E	J	C
E	222	92	55
J	83	201	126
C	98	61	210

correct: 55.1%

Table 9. Long passages

	E	J	C
E	232	85	52
J	78	162	88
C	56	92	221

correct: 57.7%

Table 10. Total (all conditions, all listeners)

	E	J	C
E	454	177	107
J	161	363	214
C	154	153	431

correct: 56.4%

Table 11. Correct Identifications

	Listeners' Own Lang.	Other Languages
Observed	463	785
Expected by chance	416	832

Table 12. Misidentifications

	Other Lang.'s as listener's own	Other lang.$_1$ as other lang.$_2$ (& vice-versa)
Observed	339	352
Expected by chance	345.5	345.5

vertical axis of Tables 1 through 10 (i.e., the labels for the rows) are the languages as presented; listed on the horizontal axis (i.e., labels for the columns) are the languages given as responses by the listeners. Thus in Table 1, for example, it can be seen that the monolingual listeners identified Japanese sentences as English 52 times. With the exception of Tables 11 and 12, the numbers along the diagonal running from top left to bottom right represent correct responses.

Discussion of Results

The overall score (number correct identifications/total responses) was 56.4% (or 58.1% if we exclude the responses of those 5 listeners who were not given the training session). This is far above the chance level of 33.3% and is therefore highly significant statistically ($p << 0.001$ via one-tailed Chi-square test on the 1 x 2 matrix consisting of correct vs. incorrect responses). The scores of those English listeners who heard the training session was higher than those who did not, 55.1% as opposed to 44.1% and this difference is highly significant. The scores for monolinguals, bilinguals, and trilinguals were 50.3%, 59.3%, and 53,2%, respectively. Of these, only the bilinguals' score is significantly better than the others. Long passages were better identified than short passages, 57.7% versus 55.1%; this difference does achieve statistical significance ($p < 0.05$). Listeners' identification of their own language is better than identification of other languages by a greater proportion than the 1:2 ratio which would be expected by chance (although this trend is reversed among the Japanese listeners) ($p << 0.01$). However, although there are slightly fewer misidentifications of other languages as listeners' own language than there are misidentifications of other language$_1$ as other language$_2$ (and vice versa), as was predicted by hypothesis (2) above, (cf. Table 12), this difference is far from achieving statistical significance and, in any case, as a trend is only true of the English listeners.

General Discussion

These results are in general agreement with previous studies and extend their validity to free conversation. Although the scores were somewhat higher than previous comparable studies, probably due to the training passage, the clear prosodic differences between the languages, and possibly, the fact that free conversation was used (thus preserving more of the "natural rhythm" of the languages), the scores were still low enough to justify characterizing the task as difficult. It is possible that the method of converting the speech signal into a "buzz" destroys certain crucial prosodic information, e.g., syllable or word boundaries, which may be necessary to evaluate any language-specific aspects of fundamental frequency variation which are synchronized with those units. Based on these promising results, however, and our own experience with this technique, we would propose that future work in this area approach the following topics:

1) The relative importance in this task of the various prosodic parameters present in the "buzz" versions of speech, i.e., fundamental frequency, amplitude, and timing, could be evaluated by holding most of them constant while one is allowed to vary -- a relatively easy thing to do if a voltage-controlled-frequency and voltage-controlled-amplitude signal generator is used, as in the present study. (Elimination of all timing information, of course, would be difficult.)

2) Would it be possible to raise listeners' scores even more by giving them a more extensive and more elaborate training session, or, possibly, by giving them feedback on the correctness of their identifications during the course of the test? If so, this might be a way of getting second language learners "attuned" to the specific prosodic patterns of the language they a are seeking to learn.

3) If such a test is performed on a large number of languages and listeners' responses presented as confusion matrices it may be possible to verify traditional linguistic prosodic classifications of languages as "syllable-timed" or "stress-timed" on the one hand and "tone language", "accent language", etc., on the other by seeing which languages are confused most often with which others under these circumstances.

Acknowledgements

We gratefully acknowledge the technical assistance of Steve Pearson, Karen Bracco, and Catherine Rodriguez-Nieto and the support of the Language Laboratory, University of California, Berkeley, and the National Science Foundation (via a grant to the Phonology Laboratory).

References

Atkinson, K. 1968. Language identification from non-segmental cues. UCLA-WPP 10:85-89.

Bonte, R. 1975. Can you identify a language by its prosody? Unpub. M.A. Thesis, University of California, Berkeley.

Bush, C. 1967. Some acoustic parameters of speech and their relationships to the perception of dialect differences. TESOL Quarterly 1:20-30.

Cohen, A. & Starkwether, J. 1961. Vocal cues to the identification of language. American J. Psychology 74:90-93.

Krones, R. 1968. Calibrating the pitch extractor. MTM, May, pp. 39-42.

Maidment, J. 1976. Voice fundamental frequency characteristics as language differentiators. Speech and Hearing Work in Progress, University College, London 2:74-93.

Richardson, J.A.C. 1973. The identification by voice of speakers belonging to two ethnic groups. Unpub. Ph.D. Dissertation, Ohio State University.

The story of [w]: An exercise in the phonetic explanation for sound patterns[1]

John Ohala and James Lorentz

Impressionistically-based, pre-theoretical taxonomies, such as many developing scientific disciplines use, are a mixed blessing. On the one hand they serve the essential, useful function of helping to organize what would otherwise be an unwieldy mass of unrelated data but on the other hand they tend to petrify the thinking of those using the taxonomies, making it difficult for them to deal with the data except insofar as they fit into the pigeonholes which have been set up. For example, many laymen, who assume a basic dichotomy between living and non-living matter, have difficulty accomodating viruses into their conceptual framework since they can exhibit properties of both of these supposedly mutually exclusive categories.

Phonology, which is still a developing science, is also plagued by this problem of unwarranted taxonomic constraints. This applies not only to traditional structural phonology but also to taxonomic generative phonology -- as has been noted before and as we will demonstrate below. This problem is dramatically evident in the treatment of [w] and other speech sounds with simultaneous and nearly equal constrictions in both the labial and velar places of articulation, e.g., [ʍ, u, kp, gb] etc., henceforth, 'labiovelars.' Simply stated, the difficulty is that many phonologists are unwilling to allow labiovelars to be classified in their phonological descriptions as both labial and velar even though their phonetic manifestation clearly has both components. For example, in most of the consonant charts accompanying descriptions of languages, [w] is placed either in the labial or velar column, not both.[1] Similarly, Chomsky and Halle (1968) argue that all labiovelars must either be primarily velars with secondary labialization or primarily labials with secondary velarization or, in their terms, must be either [+anterior] or [-anterior]. Which of the places of articulation is primary is determined by phonological evidence, not phonetic evidence. Kaisse (1975) and Anderson (1976) in more recent papers argue the same point.

Anderson, for example, reviewing the consonant inventories of some West African languages described by Ladefoged (1964) insists that doubly-articulated stops /kp/ and /gb/ have to be analyzed as labials or velars based on how they pattern in the language.[2] He suggests that since the languages Temne, Limba[3], and Sherbro have paired voiced and voiceless stops at every place of articulation except velar, where they have /k/ but not /g/, the segment /gb/ which they have must belong to the otherwise empty pigeonhole reserved for voiced velar stops. Similarly he argues that since Effutu, Kyerepong, and Nzema lack /p/, their /kp/ must fit into the

1. Paper published in Proceedings of the Third Annual Meeting of the Berkeley Linguistics Society, 1977, pp. 577-599.

voiceless labial slot. Late, Nkonya,[4] Krachi, Itsekiri, Urhobo, Idoma, and Kutep are cited as languages which have one or more labialized velar stops, /kw/, in contrast with plain velar stops, /k/, thus forcing the assignment of their double stops, /kp/, to the labial pigeonhole. Anderson further suggests that the /kp/ sound in Anum and Efik must be phonologically labial for the same reason and also because these languages lack /p/. He notes that additional evidence on the phonological identity of such double stops may be gained by observing how they pattern in phonological rules, e.g., the labiovelar stops in Kpelle and Yoruba must be velars since preceding nasals assimilating to their place of articulation show up as [ŋ] not [m].[5] In languages where labiovelars pattern as both labials and velars, e.g., the /w/ in Fula, he proposes that there must be two /w/'s, one phonologically labial and one velar.

Central to the taxonomic generative position is the not-always-clearly-defined distinction between the phonetic and phonological (or "underlying") character of speech sounds. We will comment on this in more detail below.

We will show that the question of whether labiovelars are primarily labial or velar is a pseudo-problem necessitated by the largely taxonomic approach taken by phonologists. We will do this in two ways. First, by demonstrating that the 'pigeonhole-filling' approach does not yield convincing results and, second, by showing that the kind of behavior or patterning labiovelars exhibit in phonological rules--at least those which have been cited in the literature so far--can, in general, be explained by reference to their phonetic character. It is unnecessary to posit that the phonetic character of a segment differs from its phonological or "underlying" character unless the latter terms are defined in fairly innocuous ways. Since the patterns that labiovelars exhibit are phonetically-caused, we can also show that they are universal patterns and can be found in unrelated languages throughout the world.[6] Finally, we can show that there is no contradiction in finding a single labiovelar in a given language patterning like a labial in some cases but a velar in others.

Although we will focus primarily upon the labiovelar glide [w], most of our remarks will apply with equal validity to other labiovelars, [u, kp, gb, kw, gw] etc., and we will when appropriate cite data concerning these other segments as well as data involving [w] in support of our generalizations.

The data we cite are a mixture of phonetic statements, descriptions of sound change, allophonic variation, and morphophonemic variation. These are, in fact, ultimately the same thing as far as manifestation of phonetic tendencies is concerned: today's phonetic variation is tomorrow's sound change which in turn contributes to morphophonemic variation.

The 'Pigeonhole-Filling" Criteria.

Anderson allows that the pigeonhole-filling criteria for deciding whether labiovelars are labial or velar may not decide the

issue by itself but still has 'evidential value.' However, this is probably still an overestimation of the value of this procedure.

The implication of this technique is that there is somehow a meaningful correlation between the gap in the "normal" consonant system and the extra labiovelar stop. But consonant systems lacking /p/ and /g/ are not at all uncommon even in languages outside West Africa and in languages that do not have any obvious 'leftover' segments ready to plug the holes (Sherman 1975). Furthermore, the presence of labialized velar stops /k^w/ and /g^w/ in addition to plain velar stops is also common enough in the languages of the world (see below) to be of uncertain value in helping decide how to categorize labialized velar stops in languages also having /kp/ and /gb/. Moreover, there are many West African languages which have /kp/ and/or /gb/ but no obvious gaps in the regular stop inventory that they could fit into.

If there is a correlation between such events, e.g., the absence of /p/ and presence of /kp/, there should be more languages showing the pattern that would be predicted by the product of the independent probabilities of the separate events. For example, the probability that a coin shows heads on any given toss is 1/2 = 0.5 and that a die show '6' on any given toss is 1/6 = 0.167. Thus the probability that a paired toss of a coin and a die yields heads and '6' will be 0.5 x 0.167 = 0.083. One thousand such paired tosses should yield about 83 cases of heads and '6.' Significantly more or less than that would lead us to suspect that either the coin or die (or both) were 'fixed.' In the case of segment interactions we can apply the statistical tests of significance to see whether or not the actual incidence of the combined events was disproportionately more than would be predicted by chance.

Limiting ourselves, for the sake of argument, just to the sample of 55 West African languages surveyed by Ladefoged (1964),[7] we have computed the independent probabilities of the various events of interest and give them in Table 1. From these probabilities we can predict the probabilities of combinations of these events if the combinations were due to chance. From this latter figure we can also predict how many of the 55 languages should show the combination. If the assumptions of the taxonomic phonologist are correct, there should be more languages actually showing the combined event than the results of these latter calculations would predict. As a kind of "test" of this procedure we have also computed the probability that a language from this sample will have both /kp/ and /gb/. Considering the test case first, we see that, reassuringly, there are more languages having both /kp/ and /gb/ than would be expected by chance and the difference is highly significant ($p<0.001$ by one-tailed Chi-square). In the other cases the observed frequency is either below that predicted by chance or in one case is only insignificantly more than the predicted frequency. Thus, although there may be some interaction between certain segments, e.g., /kp/ and /gb/, the evidence does not support the notion that 'gaps' in the normal stop inventory are associated with the presence of 'leftover' double-articulated stops; nor that the

Table 1. Independent probabilities of incidence of various phonological events in sample of 55 West African languages' consonant inventories (as reported by Ladefoged 1964).

Event	No. of occurrences (out of 55)	Probability of occurrence
no /p/	11	0.20
no /g/	13	0.24
/kp/	38	0.69
no /kp/	17	0.31
/gb/	29	0.53
/k^w/	19*	0.33

*includes one case of /k^f/.

Table 2. Comparison of predicted and observed incidence of various combinations of events listed in Table 1.

Combined event	Probability of occurrence due to chance	No. of languages out of 55 expected to show combined event	No. of languages actually observed	Significance if observed > expected
/kp/,/gb/	0.364	20.04	26	0.001
/kp/, no /p/	0.138	7.59	3	
/gb/, no /g/, no /kp/	0.039	2.12	3	*
/kp/,/k^w/	0.230	12.67	9	
/kp/,/k^w/, no /p/	0.046	2.53	2	

*too little data to perform test of significance.

Table 3. Incidence of labialized consonants according to place of articulation (from 706 languages catalogued by Ruhlen 1976). (Each column shows the number of languages having one or more labialized consonants at the given place of articulation.)

labial	dental	alveolar	palatal	velar	uvular	pharyngeal/ glottal
48	26	16	43	318	107	26

presence of contrasts of the sort /k/ vs. /k^w/ implies any disproportionate incidence of double stops. Moreover, the pigeonhole-filling criteria may give evidence which conflicts with the evidence from phonological rules. Thus Anderson would put the Efik /kp/ in the labial slot because it lacks /p/ and already has labialized velar /k^w/. (Actually, according to Welmers (1973) /kp/ has allophones [kp] and [p] which a taxonomist might also use as evidence of its labiality!) Nevertheless Cook (1969) reports the nasal assimilating to /kp/ sometimes appears as [ŋ].8

The problems disappear if labiovelar consonants are recognized and classified as both labial and velar and are not required to fill a single slot in the segment inventory.

Phonetic Explanations for Universal Patterns of [w].

The first universal tendency we will discuss is (A) and its corollary (A').

(A) [w] comes from both labials and velars.9

(A') labial or velar offglides, both phonemic and allophonic, are most often found on labial and velar consonants, less often on dentals and palatals.

These are fairly common patterns; evidence for (A) and (A') is given in (1)-(4):

(1) Morphophonemic alternations between /w/ and either labial or velar obstruents (or both) appear in: Nguna (Schütz 1969), Berber (Nicolas 1953), Mende (Aginsky 1935, Crosby and Ward 1944), Chichewa (Watkins 1937), Vai (Heydorn 1971b), Loma (Heydorn 1971a), Fulani (Stennes 1967), Zuni (Bunzel 1933-38), Hausa (Newman 1970), Nunggubuyu (Heath 1976), Esako (Laver 1967), Fula (Anderson 1976), Chamarro (Topping 1969). Other examples are given by Ultan (1970).

(2) There is evidence that historically labial or velar obstruents developed labial offglides or changed to /w/ in: Indo-European (Poultney 1963), Solomon Islands Melanesian (Ivens 1928, 1931).

(3) Labial offglides are predictable after the labial and velar consonants /k, g, b, w, f/ in Berber (Beni Imassen dialect, Renisio 1932).

(4) A survey of the incidence of distinctively labialized obstruents in the segment inventories of 706 languages as catalogued by Ruhlen (1976) reveals that they occur most often on velar, uvular, and labial consonants, less often on dental, alveolar, and palatal consonants; see Table 3.

The reason for this pattern can be explained first by noting that back velars, labials, and labiovelars have an important acoustic feature in common, namely, a low second formant (Lehiste and Peterson 1961, Lehiste 1964). This, of course, was the motivation for the Jakobsonian acoustic feature 'grave,' defined as 'having predominately low frequency energy,' which was applied to both lab-

ials and velars (Jakobson, Fant, and Halle 1952). Back velars and labials may be heard as labiovelar glides since their formant transitions resemble those of [w] (cf. also Liberman, Delattre, Gerstman, and Cooper 1956). The role of the acoustic similarity of these sounds in sound change has been discussed by Durand (1956) and Herbert (1975).

But why do labials and back velars produce similar acoustic effects? The reasons are known but have not received much (any?) attention in the phonological literature in spite of their clear relevance to many phonological issues. The explanation requires reference to the standing wave patterns of the resonant frequencies of the vocal tract. (We omit many details which are covered more systematically in Chiba and Kajiyama 1958, Fant 1960, Small 1973, Heinz 1974, Fidelholtz 1975.) Figure 1a represents schematically the standing wave patterns of the lowest three resonant frequencies (i.e., "formants"), R1, R2, and R3, in a uniform cylinder closed at one end and open at the other, that is, a tube resembling the vocal tract. The superimposed sine waves in the tubes represent the range of velocities of the air particles due to the standing wave patterns of the resonant frequencies. It can be seen that for all resonant frequencies there will be a velocity minimum at the glottis and a velocity maximum at the lips since the air is most constrained in its movement at the glottis but is most free to move at the lips. Additional velocity maxima and minima may be located at other places in the vocal tract for the second and higher resonances. In the case of the second resonance, which is the perceptually most salient resonance for the determination of place of articulation, it can be seen that an additional velocity minimum is located in the palatal region and an additional velocity maximum in the velar-uvular region. These are the locations of the velocity maxima and minima in a tube (vocal tract) having uniform cross-dimensional area from one end to the other. With the addition of a labial constriction, the position of the inner velocity maximum shifts forward a bit to the velar region. The locations of the velocity maxima and minima in the vocal tract under these circumstances are shown in Figure 1b. The resonant frequencies in a non-uniform tube, i.e., with one or more constrictions, can be predicted by noting whether the constriction(s) coincide with or are very near these velocity maxima and minima. The rule is: a constriction at a velocity minimum raises the resonant frequency from what it would be for a uniform tube; a constriction at a velocity maximum lowers the resonant frequency from what it would be for a uniform tube.[10]

This rule correctly predicts, for example, the high second resonance and low first resonance of the palatal vowel [i] and the high first resonance and slightly low second resonance of the pharyngeal vowel [a]. More to the point, it explains why a constriction in either the labial or back velar position will have the similar acoustic effect of lowering the second formant and why simultaneous constrictions at both labial and velar regions will lower it even more. This also explains why, of all speech

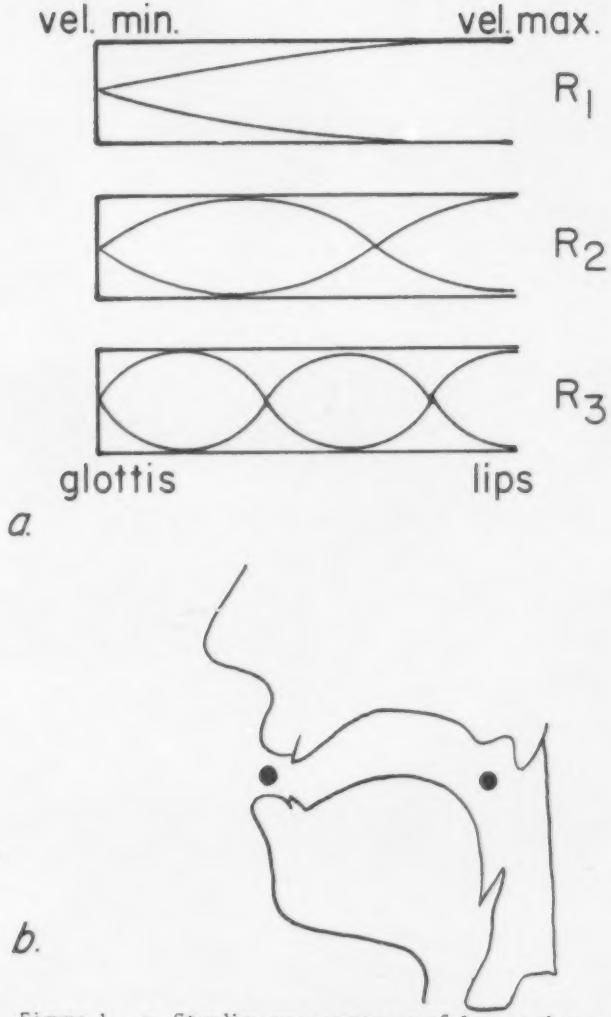


Figure 1. a. Standing wave patterns of lowest three resonant frequencies in uniform tube closed at one end and open at the other. b. Approximate location in the vocal tract of the two velocity maxima in the standing wave pattern of the second resonance.

sounds having two more or less equal places of articulation, labiovelars are so popular: they push the second formant towards an extreme low value and thus produce sounds which are auditorily very distinct from other speech segments. No other two simultaneous places of articulation (e.g., labio-palatal, palatal-velar, velar-pharyngeal, etc.) can do the same. The two simultaneous constrictions which would push the second formant to a maximally high value would be one in the palatal region and one in the pharynx immediately above the glottis. We don't find this, however, because pharyngeal constrictions are difficult and, unlike the situation with labio-velars, the two articulators, tongue tip and tongue root, are not completely independent in their movements (Lindblom, Pauli and Sundberg 1975, Lindblom 1975).

The remaining generalizations we will discuss are of particular interest since they show how [w], although both labial and velar, can, for perfectly straightforward phonetic reasons, show itself as a labial in some cases and a velar in other cases.

The second generalization regarding [w] is:

(B) When becoming a nasal or determining the place of articulation of adjacent nasals by assimilation, [w] shows itself as a velar, rarely as a labial.

Evidence for generalization (B) is found in (5) through (10) below:

(5) Nasals assimilate to [ŋ] before /w/, /kp/, and/or /gb/. For example, in Tswana (Cole 1955) m, n + w → ŋw, e.g.,
-roma "send" + wa → -roŋwa
-fena "conquer" + wa → -feŋŋwa.
Similar nasal assimilation data can be found in the following languages: Hausa (Kraft and Kraft 1973), Picuris (Trager 1971), Orizaba Nahuatal (Goller, Goller, and Waterhouse 1974), Tenango Otomi (Blight and Pike 1976), Hupa (Woodward 1964), Kpelle (Welmers 1962), Efik (Cook 1969), Ébrié (Dumestre 1970), Walapai (Redden 1966), Sierra Popoluca (Elson 1967), Maidu (Dixon 1911), Northeastern Maidu (Shipley 1956 -- but not in Shipley 1964). Also in limited environments in: Spanish (Harris 1968), Berber -- Beni Iznassen dialect -- (Renisio 1932), Mbembe (Jacquot 1962), Adioukrou (Herault 1969).

(6) Various morphophonemic, allophonic, and dialectal ŋ/w alternations. For example, in Kpelle (Welmers 1962) /w/ patterns with velars in morphophonemic alternations; e.g.:

Indefinite	Definite	
Bɔ́ɔ	ˋmɔ́ɔi	"wax"
lúu	ˋnúui	"fog, mist"
ɣîla	ˋŋilaĩ	"dog"
wée	ˋŋwéei	"white clay"

Similar data can be found in the following languages:

Iraqw languages (Tucker and Bryan 1966), Yaqui (Fraenkel 1959), Yucatec (Bowman 1959), Sinhalese (Coates and De Silva 1960), Hueyapan (Campbell 1976), Chichewa (Watkins 1937), Ngwe (Dunstan 1964), Adzera (Holzknecht 1973), Rawang (Morse 1963), Yav (Gedney 1965), Bini and Edo (Ladefoged 1964 and Westcott 1962), Mbe (Bamgbose 1967), Akan (Schachter and Fromkin 1968), Southern Paiute (Sapir 1930, Harms 1966), Yoruba (Bamgbose 1966), Kuwaa (Belleh) (Thompson 1976), Baoulé (Vogler 1968), Thonga (Passy 1914), Zoque (Wonderly 1951), Berber (Renisio 1932), Mbembe (Jacquot 1962), Mende (Crosby and Ward 1944).

Cases of historical change involving [ŋ] developing from interaction with [w] are found in (7)-(9):

(7) Ivens (1928, 1931) presents evidence from Melanesian languages that /m/ ~ /mw/ ~ /ŋg/, e.g., Ulawa /nima/; Common Melanesian /limwa/; Fiji /liŋga/ "hand."

(8) In Uto-Aztecan there is evidence of the change /w/ → /ŋ/, especially before /a/ (Munro 1973). (There is good evidence that low vowels such as /a/ are more susceptible to nasalization than higher vowels; Ohala 1975.)

(9) Additional evidence may be found in Numic (Plateau Shoshonean) (Davis 1966).

(10) The phoneme inventories of 706 languages (Ruhlen 1976) reveal the incidence of /ŋ^{w}/ outnumbering /m^{w}/ and /n^{w}/ (/ŋ^{w}/ 21 cases, /m^{w}/ 11 cases -- 9 of these from the Austro-Tai language family --, /n^{w}/ 1 case).

The explanation for (B) requires reference to the factors which create resonances and anti-resonances in the vocal tract. (Again, we simplify; for details see Fant 1960, Fujimura 1962, House 1957, Ohala 1975, Heinz 1974.) Resonances are determined by the dimensions of those airways in the vocal tract that represent a direct route from the sound source (glottis) to the point where the sound radiates to the atmosphere. Anti-resonances are determined by any airways that are cul-de-sacs branching off from this main airway. In the case of nasal consonants (see Figure 2) the direct path from the sound source (the glottis) to the radiation point (the nostrils) is via the pharyngeal and nasal airways and is substantially the same for all nasal consonants. This path is marked by filled circles in the figure where the schematic vocal tract shapes of the consonants [m], [n], [ŋ], and [w̃] are given. The main features which distinguish nasal consonants from one another, then, are differences in the oral cavity, which is a cul-de-sac branching off of the main (pharyngeal-nasal) airway and which creates anti-resonances whose frequencies depend on the dimensions of the cavity, in particular its effective length. The 'effective length,' of course, will be that length measured from the pharyngeal airway to the point of constriction in the oral cavity. In the case of multiple constrictions in the oral cavity, as in [w], it will be the back-most con-

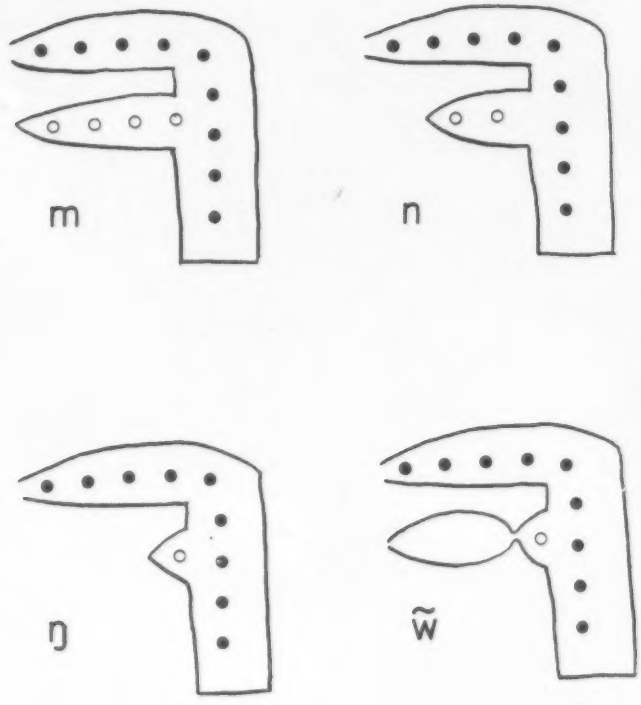


Figure 2. Schematic representation of the vocal tract shapes for the sounds [m], [n], [ŋ], and [w̃]. Filled circles indicate airways contibuting resonances of the sounds; open circles indicate airways contributing anti-resonances.

striction, provided it is small enough, which will mark the boundary of the cavity. The cavity lengths contributing the antiresonances in the vocal tract shapes in Figure 2 are marked by open circles. [w̃], then, will be most like [ŋ] rather than [m].

For similar reasons [mʲ] or the sequence [mi] is disposed to change into [n] or [ni] as happened in some Bohemian Slavic dialects (Andersen 1973), Chinese (Chen 1973), and Herero (Homburger 1949). In this case there are constrictions at both the labial and palatal regions but the effective length of the oral side cavity is determined more by the palatal constriction than by the labial constriction. This phenomenon has been well noted in acoustic phonetics (Fant 1960, Fujimura 1962) and in perceptual studies (Malécot 1956, House 1957, Gay 1970). (Of course, the consonant transitions also contribute to the auditory similarity of [mi] and [ni] and thus the sound change [pʲ] → [t] is also not uncommon; cf. Andersen 1973.)

The third generalization we offer on [w] is:

(C) When becoming a fricative or determining the place of articulation of adjacent fricatives by assimilation, [w] shows itself primarily as a labial, less often as a velar.

Evidence for generalization (C) is as follows:

(11) In Rawang (Morse 1963) [w] occurs finally and postconsonantally while [v] occurs in initial position.

(12) In Sentani (Cowan 1965) /h/ is optionally realized as [f] or [ɸ] before /w/.

(13) Similar evidence can be found in the following languages: Javanese (Horne 1961), Kirghiz (Herbert and Poppe 1963), Telegu (Lisker 1963), Selepet (McElhanon 1970), Warao (Osborn 1966), Oneida (Lounsbury 1953), Carib (Peasgood 1972), Georgian (Robins and Waterson 1952), Cashinahua-Pano (Kensinger 1963), Jeh (Gradin 1966), Chalchihuitan Tzotzil (Hopkins 1967), Cham (Blood 1967), Hungarian (Kálmán 1972), Toba Batik (Van der Tuuk 1971), Indo-European (Meillet 1964), Tenango Otomi (Blight and Pike 1976), Yolax Chinantec (Rensch 1968), Slave (Howard 1963).

(14) Pike (1943) notes that in order to make the labiovelar fricatives [ɣβ] or [xɸ],

> "the velar stricture has to be of a close variety or else its sound will be masked out and made inaudible by the vibration at the lips." [132];

cf., also Heffner (1964),

> "The fricative noises produced by the articulation of [French] [w] are slight, but such as they are, they come rather from the labial than from the velar constriction." [160]

Possible exceptions to (C) are those in (15):

(15) In Araucanian (Echeverría and Contreras 1965) [u], [w], and [γ^w] are in free variation. In Danish there is dialectal variation of [u] ~ [uu̯] ~ [ux] ~ [uk] (Andersen 1972).

With a fricativized labiovelar we are dealing with two simultaneous sources of noise produced by turbulent airflow through the labial and velar constrictions. What we must determine is why the noise produced at the labial constriction dominates. We can identify at least four reasons. The first three probably contribute to this effect but are far less essential than the fourth.

1. We have assumed thet the labial and velar constrictions in [w] were equal but there is evidence from various x-ray studies of [u], a close cousin of [w], that the labial constriction may have a slightly smaller cross-dimensional area than the velar constriction (e.g., Fant 1960). This could make the labial noise more salient than the velar noise.
2. According to Fant (1960: 274) the shape of the constriction also matters in fricative noise production. A circular constriction, which the lips can produce, is a more efficient noise generator than a slit or elliptical constriction such as would be more likely to occur at the velar region -- even if both have equal cross-dimensional areas.
3. A third factor which contributes not necessarily to making the labial noise dominant but to making the velar noise acoustically similar to the labial noise, is the fact that the air space the velar noise has to pass through, i.e., the oral cavity, shapes the noise in a way that lowers its resonant frequency towards the low center frequency of the labial noise.
4. Perhaps the most important factor is that the intensity of any sound is a function both of its inherent intensity and of the way the resonating cavities the sound passes through modify the intensity at various frequencies. The acoustic impedance seen by the velar noise source is considerably greater than that seen by the labial noise source due, in part, to the fact that the velar noise source has to pass through the narrow labial constriction whereas the labial constriction has no such constriction to attenuate its intensity.

Again we have glossed over many details, but for the reasons given the velar noise will be acoustically similar to that produced at the labial constriction and will moreover be masked by the more intense labial noise.

Central to our discussion of these last three generalizations is the assumption that if speech sounds are acoustically similar to each other, this provides the possibility of one of the sounds

changing into the other diachronically. This would happen since a listener hearing one of these auditorily ambiguous speech sounds would not know exactly how it was articulated and so when speaking himself may articulate it in a different way. This is called sound change via "acoustic imitation" by Sweet (1891) and has been further discussed by Durand (1956) and Ohala (1974a, 1974b, 1975).

A corollary of (C) is (C').

(C') Labiovelar obstruents will most likely change to labial not velar obstruents.

Evidence for (C') is the following:

(16) Indo-European k^w or kw became p regularly in Greek (with definite exceptions), in Osco-Umbrian, and in some Celtic (dialects), sporadically in Germanic (Meillet 1967, Passy 1890), e.g.,
Latin equus, Greek hippos.
Similar data can be found in the following languages: Proto-Zapotec (Suarez 1973), Songkhla (Henderson 1975), Proto-Yuman (Haas 1963). (Cf. also Campbell 1974 who lists additional examples and one counterexample.)

The fourth generalization is (D):

(D) When assimilating to adjacent vowels, it is [w]'s labial place of articulation that remains unchanged; the place of the lingual constriction may shift under the influence of the vowel's lingual configuration.

Evidence for (D) is given in (17)-(18):

(17) /w/ is realized as [w] before back, especially back rounded, vowels, but as [v], [β], or, less frequently, [ɥ], before front vowels in: Cayapa (Lindskoog and Brend 1962). Amahuaca (Osborn 1948), Jivaro (Beasley and Pike 1957), Malayalam (Sreedhar 1972), Chontal (Keller 1959), Chinese (Dow 1972), Chipaya (Olson 1967), Hawaiian (Pukui 1965), Hueyapan (Campbell 1976), Moxo (Ignaciano) (Ott and Ott 1967), Kunimaipa (Pence 1966), Gadsup (Frantz 1966), Azerbaijani (Householder 1965), Zan (Kirizia 1967), Telefol (Healey 1964), Suena (Wilson 1969b), Binandere (Wilson 1969a), Kuman and Pawaian (Trefry 1969), Mbembe (Jacquot 1962), Island Carib (Taylor 1955), Wolof (Ward 1939), Ga (Berry 1951), Quiopetec Chinantec (Robbins 1961), Kaiwa (Guarani) (Bridgeman 1961), Yareba (Weimar 1972), Piro-Arawakan (Matteson 1965), Dogon (Calame-Griaule 1965), Proto-Takanar (Girard 1971), Ulu Muar Malay (Hendon 1966), Old Irish (Cowgill 1967), Pashto (Morgenstierne and Lloyd-James 1928), Akan (Fanti) (Welmers 1946), Gã (Kropp 1968). Cf. also Tucker and Bryan (1966: 142).

(18) There is evidence in Lama (Heydorn 1971a) that p, b → v / a, ɛ, e, i, but → w / ɔ, o, u.

This is easily explained in articulatory terms. Of the two constrictions of [w], the labial and the lingual, only the lingual constriction is free to (partially) assimilate its place of articulation to that of adjacent vowels. The shift of the lingual constriction of [w] in such cases is exactly comparable to its shift in other velar consonants, [k, g, ŋ, x] etc. whose place of articulation -- as is well known -- is also influenced by neighboring vowels. The labial constriction, for obvious anatomical reasons, is not likely to assimilate to the lingual constriction of adjacent vowels.

Now we need to introduce a few qualifications regarding our data. First, the operation of (D) could lead to (i.e., trigger) the development of the pattern described in (C). If so, (C) would appear for reasons quite different from those given for it, although this would in no way lessen its purely descriptive value. Fortunately, at least some of the patterns cited in support of (C) clearly did not come about due to the action of (D).

The second qualification is to admit that some cases of the manifestation of tendencies (B) and (C) may have arisen indirectly due to the action of (A). In this case they also would not, strictly speaking, have happened for the acoustic reasons we gave above. For example, an earlier [g] would naturally cause nasals assimilating to it in place of articulation to turn up as the velar nasal [ŋ]. If this [g] later changed to [w] we would now have nasals appearing as [ŋ] before [w] but for entirely different reasons from those we gave. This sequence of events is summarized in (19).

(19) *Ng → ŋg → ŋw

Fortunately, we have included some evidence which shows that the velar nasal appears due to the physical character of [w] itself, not its velar ancestor. Likewise some apparent counterexamples to (B), that is, nasals assimilating to [w] or alternating with [w] showing up as [m], are not in fact counterexamples since the [w] is known to come from an earlier labial. Such is the case for some /w/'s in Fula (Anderson 1976), Fulani (Stennes 1967), and Hausa (Newman 1970). Thus there were different causes -- but still phonetic ones --, operating at an earlier stage in the language which determined the place of articulation of the nasal.

In (20) through (23) we give examples of languages where [w] patterns as both a labial and a velar -- the situation that supposedly could not happen if languages were constrained in the way taxonomic phonologists imagine they should be. We do not include here cases like those of Fula, Fulani, and Hausa where /w/ patterns like a labial and a velar because it derives historically from earlier labials and velars.

(20) In Tenango Otomi /h/ → [ɸ] / __ w, whereas /n/ → [ŋ] / __ k, g, h, w, ʔw, ʔV (Blight and Pike 1976).

(21) In Hueyapan /w/ is realized as [v] intervocalically after unround vowels; [v] alternates with [ŋ] finally (Campbell 1976).

(22) Spanish /w/ is realized generally in syllable initial position as [ɣ^{w}], less frequently as [β], but in allegretto speech /n/ → [ŋ] across word boundaries before /w/ (Navarro Tomas 1961, Harris 1969).

(23) Kuwaa (Belleh) word initial /w/ is occasionally realized as [ŋ^{w}] but also becomes [v] before unrounded vowels (with some exceptions) (Thompson 1976).

(24) In some dialects of Yoruba /ã/ merges with /ɔ̃/ after the labial consonants /w, kp, gb, b, f, m/; nevertheless, nasal consonants assimilating to the place of articulation of following /w/ appear as the velar [ŋ] (Ward 1952).

Conclusion.

What are the implications of the above data for the issue of whether a phonetic labio-velar is to be regarded as phonologically or "underlyingly" a labial or a velar? To answer this we must first determine what is meant by the terms 'phonologically' or 'underlyingly' in such contexts. If they simply mean that it is descriptively convenient to treat the labio-velar as a velar or a labial, that is, if the terms just have taxonomic relevance and there are no necessary empirical implications of the labelling, then we have little to argue with, except to point out that there may occasionally be instances where it would be more 'convenient' to give such sounds two labels (cf. 20-24) and that there should be no prohibition against this.

If such terms only mean that the segment was historically a velar or a labial, we also have little to quibble about except to point out that the universal physical phonetic factors we have described may override the influence of the fossilized remnants of the segment's earlier state in determining its present-day behavior (cf. 7-8).

Moreover there are many /w/'s which emerged from /u/'s which cannot themselves be identified as either labial or velar. Thus is 'underlying' means 'is derived historically from' there will be cases where the statement '/w/ is underlyingly labial (or velar)' will be meaningless or irrelevant.

If we are to interpret such statements to mean segments are psychologically velar or labial, then we seriously dispute such claims. First of all, we have shown that the behavior of /w/ (as is true of all speech sounds) is heavily influenced by physical factors. Speakers do not have to "know" the laws of physics in order for their speech to be subject to them. Second, we dispute the general operating assumption of the taxonomic generative phonologists that it is possible to discover the psychological representation of speech sounds or of grammars in general just by examining the surface sound patterns in language. Such naive psy-

chologizing has no place in a field that has aspirations of becoming a serious scientific discipline. We do not deny that there might be a psychological representation of speech sounds different in some respects from the physical phonetic realization of the the sound. But if there is, it will take some clever psychological techniques to discover it, not the application of simplistic taxonomic methods. Attempts are currently underway to develop techniques which can reveal the psychological categorization speakers apply to speech sounds (Jaeger, forthcoming). In any case, as a research strategy, we would urge that all possible phonetic explanations for sound patterns -- particularly those which are apparently universal -- be attempted before entertaining psychological explanations. We hope to have demonstrated the usefulness of this approach.

Acknowledgements.

Earlier versions of this paper were presented in the Phonology Laboratory Colloquium and before the Historical Linguistics Study Group, both at UC Berkeley. We thank the participants of these groups and also the Stanford Phonology Archive for helpful comments and data. This research was supported in part by NSF.

Footnotes.

1. The traditional consonant chart of the International Phonetic Association is an exception to this generalization.
2. In this, Anderson follows Ladefoged's lead.
3. However, Ladefoged states 'k occurs word initial and after ŋ; g or ɣ occur elsewhere.' It is thus possible that in some environments [g] and [gb] would contrast and remove the motivation to put /gb/ in the voiced velar slot.
4. However, Reineke (1972) reports that it is not clear that [k^w] represents a separate phoneme in Nkonya distinct from /k/ since it sometimes alternates with [ko]. In any event, she also reports that n, m → ŋ/ $\underline{\quad}$ kp, w.
5. We have been unable to verify this statement. All the sources we have consulted, Welmers (1962) for Kpelle and Ward (1952), Bamgbose (1966), Stevick and Aremu (1963) for Yoruba, report that the nasal before the labio-velar stops in these languages is [ŋm] not [ŋ] as claimed by Anderson.
6. And therefore that the patterns are not due to language-specific factors.
7. We have also accepted the consonant inventories as presented by Ladefoged and have made no attempt to correct probable errors. For example, Ladefoged does not list a /g/ for Mende but two other sources, Aginsky (1935) and Crosby and Ward (1944), do.

8. Cf. also note 4.
9. There are, of course, other sources of [w], e.g., back rounded vowels [u] and [o], [ɫ] (velarized lateral), and the vocalic transitions between such vowel sequences as /oe/.
10. See Fant (1960: 86-7) for the physical reasons motivating this rule.

References.

Abbreviations Used.

A&U	Afrika und Ubersee.
AL	Anthropological Linguistics.
AUA	Annales de l'Universite d'Abidjan. Serie H.
BSO(A)S	Bulletin of the School of Oriental (and African) Studies, London.
IJAL	International Journal of American Linguistics.
JAL	Journal of African Languages.
JASA	Journal of the Acoustical Society of America.
JWAL	Journal of West African Languages.
Lg	Language.
PL	Pacific Linguistics.
SIL	Summer Institute of Linguistics.
UCPL	University of California Publications in Linguistics.
WPLU	Working Papers in Language Universals, Stanford.

Aginsky, E.G. 1935. A grammar of the Mende Language. Lg. Diss. 20.

Andersen, H. 1972. Diphthongization. Lg. 48. 11-50.

_____. 1973. Abductive and deductive change. Lg. 49. 765-793.

Anderson, S.R. 1976. On the description of multiply-articulated consonants. J. of Phonetics. 4. 17-27.

Bamgbose, A. 1966. A grammar of Yoruba. Cambridge Univ. Press.

_____. 1967. Notes on the phonology of Mbe. JWAL. 4. 5-11.

Beasley, D. and K.L. Pike. 1957. Notes on Huambisa phonemics. Lingua Posnaniensis. 6. 1-8.

Berry, J. 1951 The pronunciation of Ga. Cambridge: Heffer.

Blight, R.C. and E.V. Pike. 1976. The phonology of Tenango Otomi. IJAL. 42. 51-7.

Blood, D.L. 1967. Phonological units in Cham. AL. 9.8. 15-32.

Bowman, E. 1959. An attempt at an analysis of Modern Yucatec from a small corpus of recorded speech. AL. 1.4. 43-86.

Bridgeman, L.I. 1961. Kaiwa (Guarani) phonology. IJAL. 27. 329-34.

Bunzel, R.L. 1933-38. Zuni. In: F. Boas (ed.). Handbook of American Indian Lnaguages, Part 3. Columbia Univ. Press. 396-515.

Calame-Grianle, G. 1965. Ethnologie et langage. La parole chez lez Dogon. Gallimard.

Campbell, L. 1974. Phonological features: problems and proposals. Lg. 50. 121-36.

Campbell, R. Joc. 1976. Underlying /η^{w}/ in Hueyapan Nahuatal. IJAL. 42.1. 46-50.

Chen, M. 1973. Nasals and nasalization in the history of Chinese. Doc. diss., Univ. of California, Berkeley.
Chiba, T. and M. Kajiyama. 1958. The vowel: its nature and structure. Tokyo: Kaiseikan.
Coates, W.A. and M.W.S. de Silva. 1960. The segmental phonemes of Sinhalese. Univ. of Ceylon Review. 18. 163-75.
Cole, D.T. 1955. An introduction to Tswana grammar. London: Longmans, Green, & Co.
Cook, T.L. 1969. The pronunciation of Efik for speakers of English. Bloomington: African Studies Program.
Cowan, H.K.J. 1965. Grammar of the Sentani language. The Hague: Martinus Nijhoff.
Cowgill, W. 1967. On the fate of *w in Old Irish. Lg. 43. 129-38.
Crosby, K.H. and I.C. Ward. 1944. An introduction to the study of Mende. Cambridge: Heffer.
Davis, I. 1966. Numic consonantal correspondences. IJAL. 32.2. 124-40.
Dixon, R.B. 1911. Maidu. In: F. Boas (ed.). Handbook of American Indian Languages, Part 1. Govt. Printing Office. 679-734.
Dow, F.D.M. 1972. An outline of Mandarin phonetics. (Oriental Monograph Series, 10) Canberra: Australian National University, Faculty of Asian Studies.
Dumestre, G. 1970. Contribution a l'étude phonologique de l'ébrié (Basse Cote-d'Ivoire). AUA. 3. 19-29.
Dunstan, E. 1964. Towards a phonology of Ngwe. JWAL. 1. 39-42.
Durand, M. 1956. Du rôle de l'auditeur dans la formation des sons du langage. J. Psych. Norm. et Path. 52. 347-55.
Echeverría, M.S. and H. Contreras. 1965. Araucanian phonemics. IJAL. 31. 132-5.
Elson, B.F. 1967. Sierra Popoluca. In: McQuown, N. (ed.). Linguistics. [Vol. 5 of Handbook of Middle American Indians.] Austin: U. Texas Press. 269-90.
Fant, G. 1960. Acoustic theory of speech production. The Hague: Mouton.
Fidelholtz, J.L. 1975. An introduction to acoustics for linguistics. Ms.
Fraenkel, G. 1959. Yaqui phonemics. AL. 1.5. 7-18.
Frantz, C.I. and M.E. Frantz. 1966. Gadsup phoneme and toneme units. PL. 7. 1-11.
Fujimura, O. 1962. Analysis of nasal consonants. JASA. 34. 1865-75.
Gay, T. 1970. Effect of filtering and vowel environment on consonant perception. JASA. 48. 993-8.
Gedney, W.J. 1965. Yay, a Northern Tai language in North Vietnam. Lingua. 14. 180-93.
Girard, V. 1971. Proto-Takanan phonology. UCPL 70.
Goller, T.R., P.L. Goller, and V.G. Waterhouse. 1974. The phonemes of Orizaba Nahuatal. IJAL. 40. 126-31.
Gradin, D. 1966. Consonantal tone in Jeh phonemics. Linguistic Circle of Saigon Publications. 3. 41-53.

Harms, R.T. 1966. Stress, voice, and length in Southern Paiute. IJAL. 32.3. 228-35.
Harris, J.W. 1969. Spanish phonology. Cambridge, Mass.: MIT Press.
Healey, A. 1964. Telefol phonology. PL. 3.
Heath, J. 1976. Linguistic diffusion in Arnhem Land. Australian Institute of Aboriginal Studies.
Heffner, R-M.S. 1964. General phonetics. Madison: Univ. Wisconsin Press.
Heinz, J.M. 1974. Speech acoustics. In: T.A. Sebeok (ed.). Current trends in linguistics, Vol. 12: Linguistics and adjacent arts and sciences, Part 10: Phonetics. The Hague: Mouton. 2241-80.
Henderson, E.J.A. 1975. Feature shuffling in Southeast Asian languages: or how big is a segment? Paper presented at the Sino-Tibetan Conference, Berkeley.
Hendon, R.S. 1966. The phonology and morphology of Ulu Muar Malay. Yale Univ. Pub. in Anthropology. 70.
Herault, G. 1969. Étude phonétique et phonologique de l'adioukrou. AUA. 2.
Herbert, R.K. 1975. Reanalyzing prenasalized consonants. Studies in African Linguistics. 6. 105-23.
Herbert, R.J. and N. Poppe. 1963. Kirghiz manual. Indiana Univ. Publications, Uralic and Altaic Series, Vol. 33. Bloomington: Indiana Univ.
Heydorn, R.W. 1971a. Grammar of the Lɔma language with some Kɔnɔ words in the vocabulary. AU. 54. 77-99.
_____. 1971b. The Vai language in Liberia. AU. 54. 149-205.
Holzknecht, K.G. 1973. The phonemes of the Adzera language. PL. 38. 1-11.
Homburger, L. 1949. The Negro-African languages. London: Routledge & Kegan Paul.
Hopkins, N.A. 1967. A short sketch of Chalchihuitan Tzotzil. AL. 9.4. 9-25.
Horne, E.C. 1961. Beginning Javanese. New Haven: Yale Univ. Press.
House, A.S. 1957. Analog studies of nasal consonants. J. Speech & Hearing Disorders. 22. 190-204.
Householder, Fred W. 1965. Basic course in Azerbaijani. The Hague: Mouton.
Howard, P.G. 1963. A preliminary preservation of Slave phonemes. Studies in the Athapaskan languages, by Harry Hoijer and others. UCPL 29. 42-7.
Ivens, W.G. 1928. A grammar of the Lau language, North East Coast of Big Mala, Solomon Islands. BSOS. 5. 323-43.
_____. 1931. A grammar of the language of Kwara 'Ae, North Mala, Solomon Islands. BSOS. 6. 679-700.
Jacquot, A. 1962. Notes sur la phonologique du beembe (Congo). JAL. 1. 232-42.
Jaeger, J.J. Forthcoming. The psychological reality of phonemes revisited.

Jakobson, R., C.G.M. Fant, and M. Halle. 1952. Preliminaries to speech analysis. Cambridge, Mass.: MIT Press.
Kaisse, E. 1975. Segments with internal structure -- the Greek labiovelars. Papers from the Annual Meeting of the Northeast Linguistics Society. 5. 144-52.
Kálmán, B. 1972. Hungarian historical phonology. In: L. Benko and S. Imre (eds.). The Hungarian language. The Hague: Mouton. 49-83.
Keller, K.C. 1959. The phonemes of Chontal (Mayan). IJAL. 25. 44-53.
Kensinger, K.M. 1963. The phonological hierarchy of Cashinahua (Pano). Studies in Peruvian Indian Linguistics: I. Norman: SIL. 9. 207-17.
Kiziria, A.I. 1967. Zanskij jazyk Jazyki narodov SSSR, Vol. 4. 62-76.
Kraft, C.H. and M.G. Kraft. 1973. Introductory Hausa. Berkeley: Univ. of California Press.
Kropp, M.E. 1968. An analysis of the consonant system of Gã. JWAL. 5.1. 59-61.
Ladefoged, P. 1964. A phonetic study of West African languages. Cambridge Univ. Press.
Laver, J. 1967. A preliminary phonology of the Ayele dialect of Esako. JWAL. 4.2. 53-6.
Lehiste, I. 1964. Acoustical characteristics of selected English consonants. IJAL. Publication 34.
Lehiste, I. and G.E. Peterson. 1961. Transitions, glides, and diphthongs. JASA. 33. 268-77.
Liberman, A.M., P.C. Delattre, L.J. Gerstman, and F.S. Cooper. 1956. Tempo of frequency change as a cue for distinguishing classes of speech sounds. J. Experimental Psychology. 52. 127-37.
Lindblom, B. 1975. Experiments in sound structure. Plenary address, 8th Int'l. Congress of Phonetic Sciences, Leeds, 17-23 August 1975.
Lindblom, B.E.F., S. Pauli, and J. Sundberg. 1975. Modelling co-articulation in apical stops. In: G. Fant (ed.). Speech communication. Proceedings of the Speech Communication Seminar, Stockholm, 1-3 August 1974. Vol. 2: Speech production and synthesis by rule. Stockholm: Almqvist and Wiksell. 87-94.
Lindskoog, J.N. and R.M. Brend. 1962. Cayapa phonemics. Studies in Ecuadorian Indian Languages: I. B. Elson (ed.). Norman: SIL. 7. 31-44.
Lisker, L. 1963. Introduction to spoken Telegu. New York: American Council of Learned Societies.
Lounsbury, F.G. 1953. Oneida verb morphology. New Haven: Yale Univ. Press. Yale Univ. Publications in Anthropology. 48.
Malécot, A. 1956. Acoustic cues for nasal consonants: an experimental study involving a tape-splicing technique. Lg. 32. 274-84.
Matteson, E. 1965. The Piro (Arawakan) language. Berkeley: Univ. of California Press. UCPL. 42.
McElhanon, K.A. 1970. Selepet phonology. PL. 14.

Meillet, A. 1964. Introduction a l'étude comparative des langues indo-europeennes. Univ. of Alabama Press.
_____. 1967. The comparative method in historical linguistics. Paris: Librairie Honore Champion, Editeur.
Morgenstierne, G. and A. Lloyd-James. 1928. Notes on the pronunciation of Pashto. BSOS. 5. 53-62.
Morse, R.H. 1963. Phonology of Rawang. AL. 5.5. 17-41.
Munro, P. 1973. Proto-Uto-Aztecan *w -- one source for Luiseño ŋ. IJAL. 39. 135-6.
Navarro Tomas, T. 1961. Manual de pronunciacion Española. Publicaciones de la revista de filologia Española. Num. 3. Madrid.
Newman, P. 1970. Historical sound laws in Hausa and in Dera (Kanakuru). JWAL. 7. 39-51.
Nicolas, F. 1953. La langue berbère de Mauritanie. Mémoires de l'Institut Francais d'Afrique Noire, Dakar. 33.
Ohala, J.J. 1974a. Phonetic explanation in phonology. Papers from the Natural Phonology Parasession. Chicago Linguistics Society. 251-74.
_____. 1974b. Experimental historical phonology. Historical Linguistics II. Theory and description in phonology. Amsterdam: North Holland. 353-89.
_____. 1975. Phonetic explanation for nasal sound patterns. In: C.A. Ferguson, L.M. Hyman, and J.J. Ohala (eds.). Nasalfest, Papers from a Symposium on Nasals and Nasalization. Stanford: Language Universals Project. 289-316.
Olson, R.D. 1967. The syllable in Chipaya. IJAL. 33. 300-4.
Osborn, H. 1948. Amahuaca phonemes. IJAL. 14. 188-90.
Osborn, H.A. Jr. 1966. Warao I: phonology and morphophonemics. IJAL. 32.2. 108-23.
Ott, W.G. and R.H. Ott. 1967. Phonemes of the Ignaciano language. Linguistics. 35. 56-60.
Passy, P. 1890. Étude sur les changements phonétiques. Paris: Librairie Firmin-Didot.
_____. 1914. La langue thonga. Miscellanea Phonetica. 1. 27-32.
Peasgood, E.T. 1972. Carib phonology. Languages of the Guianas. J.E. Grimes (ed.). Norman: SIL. 35. 35-41.
Pence, Alan. 1966. Kunimaipa phonology: hierarchical levels. PL. 7. 49-67.
Pike, K.L. 1943. Phonetics. Univ. of Michigan Press.
Poultney, J.W. 1963. Evidence for Indo-European alternation of Initial /g^w/ and /w/. Lg. 39. 398-408.
Pukui, Mary K. and S.H. Elbert. 1965. Hawaiian-English dictionary. Honolulu: Univ. of Hawaii Press.
Redden, J.E. 1966. Walapai I: phonology. IJAL. 32. 1-16.
Reinecke, B. 1972. The structure of the Nkonya language. Leipzig: VEB Verlag Enzyklopädie.
Renisio, A. 1932. Étude sur les dialectes berbères. Paris: Editions Ernest Leroux.

Rensch, C.R. 1968. Proto Chinantec phonology. Papeles de la Chinantla. 6.
Robbins, F.E. 1961. Quiotepec Chinantec syllable patterning. IJAL. 27. 237-50.
Robins, R.H. and N. Waterson. 1952. Notes on the phonetics of the Georgian word. BSOS. XIV. 1. 55-72.
Ruhlen, M. 1976. A guide to the languages of the world. Stanford.
Sapir, E. 1930. Southern Paiute, a Shoshonean language. Proc. Am. Acad. of Arts and Sci. 65.1.
Schachter, P. and V. Fromkin. 1968. A phonology of Akan: Akuapen, Asante, and Fante. Working Papers in Phonetics (UCLA) 9.
Schütz, A.J. 1968. A pattern of morphophonemic alternation in Nguna New Hebrides. PL. 15. 41-7.
Sherman, D. 1975. Stops and fricative systems: a discussion of paradigmatic gaps and the question of language sampling. Working Papers on Language Universals (Stanford) 17. 1-31.
Shipley, W. 1956. The phonemes of Northeastern Maidu. IJAL. 22. 233-7.
Small, A.M. 1973. Acoustics. In: F.D. Minifie, T.J. Hixon, and F. Williams (eds.). Normal aspects of speech, hearing, and language. Englewood Cliffs: Prentice-Hall. 11-72.
Sreedhar, M.V. 1972. Phonology of the Cochin dialect of Malayalam. Int'l. J. of Dravidian Linguistics. 1. 100-25.
Stennes, L.H. 1967. A reference grammar of Adamawa Fulani. African Languages Monograph. Mich. State Univ. 8.
Stevick, E.W. and O. Aremu. 1963. Yoruba basic course. Washington, D.C.: Foreign Service Inst.
Suarez, J.A. 1973. On Proto-Zapotec phonology. IJAL. 39. 236-49.
Sweet, H. 1891. A new English grammar. Oxford: Clarendon Press.
Taylor, D. 1955. Phonemes of the Hopkins (British Honduras) dialect of Island Carib. IJAL. 21. 233-41.
Thompson, R.B. 1976. A phonology of Kuwaa (Belleh). M.A. thesis. San Jose State Univ.
Topping, D.M. 1969. A restatement of Chamorro phonology. AL. 11.2. 62-78.
Trager, F.H. 1971. The phonology of Picuris. IJAL. 37. 29-33.
Trefry, D. 1969. A comparative study of Kuman and Pawaian. PL. Monograph 13.
Tucker, A.N. and M. Bryan. 1966. Linguistic analyses of the non-Bantu languages of North-Eastern Africa. Oxford Univ. Press.
Tuuk, H.N. Van der. 1971. A grammar of Toba Batak. The Hague: Martinus Nijhoff.
Ultan, R. 1970. Some sources of consonant gradation. Working Papers on Lnaguage Universals (Stanford) 2. C1-C30.
Vogler, P. 1968. Esquisse d'une phonologie du baoulé. AUA. 1. 5-17.
Ward, I.C. 1939. A short phonetic study of Wolof (Jolof) as spoken in Senegal. Africa. 12. 320-34.
_____. 1952. An introduction to the Yoruba language. Cambridge: Heffer.

Watkins, M.H. 1937. A grammar of Chichewa. Lg. Diss. 24.
Weimer, H. and N. Weimer. 1972. Yareba phonemes. Te Reo. 15. 52-7.
Welmers, W.E. 1946. A descriptive grammar of Fanti. Lg. Diss. 39.
_____. 1962. The phonology of Kpelle. JAL. 1. 69-93.
_____. 1973. African language structures. Berkeley: Univ. of California Press.
Westcott, R.W. 1962. A Bini grammar. Part I: phonology. African Languages and Area Center. Mich. State Univ.
Wilson, D. 1969a. The Binandere language family. PL. 18. 65-86.
_____. 1969b. Suena phonology. PL. 18. 87-93.
Wonderly, W.L. 1951. Zoque II: phonemes and morphophonemics. IJAL. 17. 105-23.
Woodward, M.F. 1964. Hupa phonemics. Studies in California Linguistics. William Bright (ed.). Berkeley: Univ. of California Press. UCPL. 34. 199-216.

Testing Hypotheses Regarding the Psychological Manifestation of Morpheme Structure Constraints*

John J. Ohala and Maniari Ohala+

Linguists have for many decades recognised that not all possible sequences or types of sounds occur in the morphemes of any single language. These restrictions were called 'phonotactic' or 'distributional' constraints by the structuralists and 'morpheme structure' rules or conditions (MSC's) by generative phonologists. Although these constraints may originally have been conceived of as merely descriptive statements one could make about the sound patterns of the language, there is sufficient evidence now that native speakers are also aware of them, i.e., that many of the constraints noted by the linguist are psychologically real (Esper 1925, Greenberg and Jenkins 1964, Zimmer 1969, Bronkin 1971, M. Ohala 1972, Wright 1975.)

It is of interest, then, to try to find out how these morpheme structure constraints are represented in speakers' heads and how they use them. This paper is a progress report on our attempt to do this.

There have been various proposals about the nature of MSC's, although often the proposals are not perfectly explicit. Thus in reviewing some of these claims we hope we will be forgiven if we read into a certain view details that were not intended. Also it is not always clear how much psychological reality various authors intend their proposals to have. In any case our primary purpose is not to do battle with straw men but to attempt to pit one reasonably explicit and superficially plausible hypothesis regarding MSC's against another.

Greenberg and Jenkins (1964) (hereafter G&J) offer one of the most explicit statements regarding the psychological manifestation of MSC's: we will call it the phoneme substitution procedure. According to them the degree of closeness of a given word to the native pattern would be inversely proportional to the number of zero-, one-, two-, and up to n-phoneme substitutions for the n original phonemes in the word which resulted in existing words in the language, as is given in Table 1.

To find the distance from English of klæb, klɛb, and zbyk using their phoneme substitution procedure, we first note that a zero phoneme substitution would not yield an existing word as this would happen only if the word we start with were already an existing word, which none of these are. We next try changing single phonemes. If we change the first phoneme we can get slab from klæb, but cannot derive any existing words from klɛb or zbyk.[1]

---

* Revised version of a paper first presented at the Annual Meeting of the Linguistic Society of America, San Francisco, 28-30 December 1975.

+ Linguistics Program, San Jose State University, San Jose, Calif.

Table 1. Example of how Greenberg & Jenkins' phoneme-substitution method is used to determine "distance" from English of a given word.

Type of Substitution \ Word	[klæb]	[klɛb]	[zbyk]
CCVC	-	-	-
CCVC	slab	-	-
CCVC	crab	-	-
CCVC	club	club	-
CCVC	clam	Clem	-
CCVC	stab	-	-
CCVC	crib	crib	-
CCVC	clip	clip	-
CCVC	slob	slob	-
CCVC	slam	phlegm	-
CCVC	cram	Kress	-
CCVC	grub	grub	creek
CCVC	cream	cream	-
CCVC	spat	bread	-
CCVC	flip	flip	-
CCVC	gruff	gruff	gruff
Total successful Substitutions	15	12	2
Difference between total and 17 = predicted "distance" from English	2	5	15

Changing the second phoneme also succeeds with klæb but not the other two. And so on, for two, three, and finally four phoneme substitutions, the latter type, of course, yielding an existing word for any sequence of four phonemes. We then total the number of substitutions yielding an existing word and subtract that from the number of possible substitutions plus one; this is 17 in the case of four phoneme sequences. Thus klæb, klɛb and zbyk are predicted to be respectively 2, 5, and 15 units away from English on a 16 point scale. The only substitutions G&J allow in this procedure are consonant substitutions for consonants and vowels for vowels. They rule out vowel for consonant subsititutions or vice-versa, and they rule out phoneme additions or deletions.

Proposals regarding MSC's in early generative phonology (Halle 1959, 1962, Stanley 1967) (hereafter EGP) would seem to include the following claims:

a. That the function of the MSC's was to differentiate between existing or admissible phoneme sequences and inadmissable sequences, that is, a binary classification -- or under some interpretations, a ternary classification between existing words, accidental gaps, and inadmissable forms.

b. That the MSC's are stated in terms of features, not phonemes.

c. That the MSC's are separate from the lexical entries themselves and they would presumably be relatively small in number in comparison to the number of lexical entries. That is, they would just capture the rather broad generalizations such as, that in English a nasal consonant may be the second member of an initial cluster only if [s] is the first, but not including such specific constraints as that [ɛ] may not appear after the initial cluster [kl] if [b] follows it.

In The Sound Pattern of English Chomsky and Halle (1968) (hereafter C&H) offer a new, rather explicit proposal regarding the form and function of MSC's. They claim that the degree of closeness to the native pattern of a given form would be inversely proportional to the complexity of the rule -- measured by feature counting -- which would differentiate the form from all the entries in the regular lexicon. That is, we find the simplest rule, stated in features, which would not change any entry in the regular lexicon but would change an irregular or non-existing form. We then count the number of features utilized in this rule (which presumably need not incorporate any features which can be specified by universal marking conventions) and then take the distance from the regular pattern to be the inverse of this number. Some examples are given in Table 2. 2a gives the rule offered by C&H as being the simplest one which differentiates blik from all entries in the English lexicon. Since this form is at best an accidental gap in the English lexicon the rule must be very specific and thus very complicated if it is to affect blik but not affect any regular English morpheme. The rule takes 17 features and thus is $\frac{1}{17}$ or 0.059 units away from English on a scale that presumably runs from zero to 0.5. 2b gives the much simpler rule which differentiates the famous bnik from the English lexicon, and 2c and 2d give the rules which we offer as differentiating

the made-up forms xlox and xrit, and mløf and spøf, respectively, from the regular forms in the English lexicon.

There are then a variety of quite opposing claims made by these various models. To mention a few:

Both the G&J and the C&H models would predict that speakers are capable of making scalar judgements on the degree to which a given form adheres to the pattern of their native language. The EGP model however would predict that speakers could make only binary or possibly only ternary distinctions. G&J's experiments proved that subjects are capable of scalar judgements.[2]

G&J posit that in making their judgements speakers refer to the existing words in their lexicon; except for the phoneme substitution procedure itself, they need not store lexicon-independent rules. The EGP model seems to posit that the speakers refer to independent rules -- the MSC's.

Related to these two opposite predictions is that alluded to previously. That is, whether the speaker can detect deviations from only the most general patterns of the language or whether he can detect even the most specific deviations. The EGP model would presumably allow only the broad generalizations to be reflected in MSC's, whereas the G&J model, since it has the test forms compared to the entire list of existing morphemes, would permit deviations from even the most specific patterns to be detected. It was this difference between the G&J and the EGP model that we sought to test first. To do this, we attempted to find pairs of English words such that each model would presumably give different predictions about the distance from English of the two members of each pair. Table 3 gives five such word pairs. In the case of the first pair klɛb and klæb, we have already explained why the G&J model would predict klɛb to be judged further from English than klæb (refer to Table 1). We would presume that the EGP model would fail to incorporate an MSC that differentiates them. The other four word pairs, θlɛz, θlɛd; ʃriz, ʃrɪd; flʊt, flɪg; sθɛʃ, sθɪp would pattern like klɛb, klæb as far as the two models' predictions are concerned.

Then we had to obtain native English speakers' subjective judgements on the "distance" of these words from the English pattern. Before describing how this was done, however, let us mention the second point we tested: namely, whether the decision on the regularity or lack of it of a given form is based on only part of the form or on the whole form. The C&H model, especially in the case of words violating very fundamental patterns, would allow only the most serious violation, and then only one instance of it, to influence speakers' reaction to it. For example, as given in 2d, even though there are two violations of English MSC's in mløf, namely a bad initial cluster and a non-English segment [ø], as opposed to only one violation in spøf, namely, the same non-English segment, according to the C&H model, both would be equally far from the English pattern since they would both be differentiated from English by the same minimal rule -- the one given in 2d. The G&J model, on the other hand, predicts that all parts of the words will affect its estimated degree of deviation from the regular lexicon. Again, we attempted to test

Table 2

(a) Minimal rule which differentiates blik from English lexicon. (C&H, SPE: 417)

$$[+\text{cons}] \rightarrow [-\text{lateral}] \;/\; [-\text{seg}] \begin{bmatrix} +\text{cons} \\ -\text{voc} \\ +\text{ant} \\ -\text{cor} \\ +\text{voice} \end{bmatrix} ____ \begin{bmatrix} +\text{high} \\ -\text{back} \\ -\text{tense} \end{bmatrix} \begin{bmatrix} +\text{cons} \\ -\text{voc} \\ -\text{ant} \\ -\text{cor} \\ -\text{cont} \\ -\text{voice} \end{bmatrix}$$

Number of features required: 17
Distance from English: $\frac{1}{17}$ = 0.059 (minimum: 0; maximum: 0.5)

(b) Minimal rule which differentiates bnik from English lexicon. (C&H, SPE: 417)

$$[+\text{cons}] \rightarrow [-\text{nasal}] \;/\; [-\text{seg}] \begin{bmatrix} +\text{cons} \\ -\text{cont} \end{bmatrix} ____$$

Number of features required: 5
Distance from English: $\frac{1}{5}$ = 0.2

(c) Minimal rule which differentiates xlox and xrit from English lexicon.

$$\begin{bmatrix} +\text{cons} \\ +\text{cont} \end{bmatrix} \rightarrow [-\text{back}]$$

Number of features required: 3
Distance from English: $\frac{1}{3}$ = 0.33

(d) Minimal rule which differentiates mløf and spøf from English lexicon.

$$\begin{bmatrix} +\text{syllabic} \\ -\text{back} \end{bmatrix} \rightarrow [-\text{round}]$$

which of these opposing claims was closer to the truth by getting native English speakers' jugdements on word pairs such as mlɵf and spɵf and the other three pairs given in Table 4. In the case of the pair sfʊb, sfɛt, we are assuming -- perhaps wrongly -- that C&H would count the initial sf- cluster as a deviation from the regular English pattern.

Our method of obtaining native English speakers' subjective judgements regarding the degree of Englishness or non-Englishness of such words was similar to that of G&J. All of these test words and some additional words included to fill out the 16 point G&J scale, were spoken by a phonetician in two different randomized orders and recorded on tape using a high fidelity recording system. The tape was played to individual subjects (either over a loudspeaker in a sound-treated room or over high quality earphones) and they were asked to judge the "distance" of each word from English on an 11 point scale, and to mark their responses on an answer sheet provided to them.[3] Instructions given ad lib, to the subjects, went approximately as follows: "You will hear a list of words. Some will be English words, some have been obtained from languages close to English and some from languages quite unlike English. We want you to listen to each word and to judge by how it sounds how far it is from the pattern of English words, using an 11 point scale. If it is an English word check '1' off on the answer sheet; if it is very far away from English mark '11' on the answer sheet. The closer to English you judge a word to be the smaller should be the number you mark; the further away the word is from English, the larger should be the number you mark. Please try to use the entire 11 point scale." Then a minute or so of the tape was played to familiarize the subjects with the format of the tape and to hear the type of words used. Then any questions they had were answered, and the tape was started again from the beginning and the subjects gave their judgements on the words. Before the words were played to them in the second randomized order the subjects were told that since they now had a better idea of the total range of variation of the word types they could, on this second run, change their ratings of the words if they liked.

The subjects were adult speakers of English and were non-linguists. We obtained judgements from 21 subjects on those 5 pairs designed to test the different predictions of the G&J and the EGP models. For the testing of the G&J versus the C&H model we obtained ratings from 11 subjects.

Some subjects failed to use the entire 11 point scale. To make their responses comparable to the others their ratings were normalized, i.e., their scale was "stretched" linearly such that the extreme scores still equalled 1 and 11.

In tabulating the results it was the difference between the ratings given to the two words of a pair on any single run that was counted. For example if one subject rated word A 5 and word B 7, this would be counted the same as another subject who gave the ratings 2 and 4, respectively, to the the same words, since in both cases their ratings of the two words differed by 2, or, more specifically, by +2, as we must keep track of the sign of the difference. As we were simply interested in whether or not one member of a

Table 3

WORD PAIR	PREDICTIONS		RESULTS		
	According to Greenberg & Jenkins model	According to Early Gen. Phonol. model	Order	Mean diff. (on 11 pt. scale)	Test of significance (one-tailed t-test)
A B					
klɛb/klæb	A > B	A = B	A > B	0.276	
θlɛz/θlɛd	A > B	A = B	A > B	1.557	
ʃriz/ʃrɪd	A > B	A = B	B > A	0.107	
flʊt/flɪg	A > B	A = B	B > A	0.550	
sθɛʃ/sθip	A > B	A = B	A > B	1.483	
Average of All 5 pairs:				0.532	p >> 0.001

Table 4

WORD PAIR	PREDICTIONS		RESULTS		
	According to C & H	According to G & J model	Order	Mean diff. (on 11 pt. scale)	Test of significance (one-tailed t-test)
A B					
mlɔʒ/mlit	A = B	A > B	A > B	0.514	
mløf/spøf	A = B	A > B	A > B	3.014	
xlox/xrit	A = B	A > B	A > B	2.209	
sfʊb/sfɛt	A = B	A > B	A > B	2.018	
Average of All 4 pairs:				1.914	p >> 0.001

pair was judged further from English than the other, a one-tailed t-test was performed to test the significance of the difference that was found in the predicted direction -- that is, whether the mean difference in scores was significantly distinct from zero.

The results are given in Tables 3 and 4. To consider 3 first, it can be seen in two of the 5 pairs, namely, ʃriz/ʃrɪd, and flʊt/flɪg that the difference in scores was small and in the opposite direction of that predicted by G&J. In the other three pairs the difference was in the direction they predicted but in only 2 of these was the difference very large. The average difference for all 5 pairs was 0.532, which is highly significant. It may be safe to say, then, that deviations from rather narrow, quite particular patterns in the language, contribute to the 'distance' of the word from the native pattern, as G&J predict.[4]

The results of the other test are given in the right-most columns of Table 4. In all four word pairs, the difference in ratings is in the direction predicted by G&J and in three of these cases the difference is very large. The average difference for the 4 pairs was 1.914, which is highly significant. This seems to establish fairly conclusively that the whole word and not just part of it contributes to speakers' ratings of it. We believe the results of both studies also strongly imply (although perhaps do not conclusively prove) that the speakers make reference to the words in their lexicon, not to lexicon-independent rules, to make judgements in tests such as these. If they used rules, these results suggest they would have to incorporate implausibly fine detail on the sound patterns of the language. In general then, our results give the nod to the G&J model.[5]

There are still many other details to test, of course, namely whether the speaker operates with phonemes or features in comparing the "test" word with existing words and whether or not, as G&J somewhat arbitrarily maintain, that in seeing how much you had to modify a word to make it into an existing word, deletions or additions are not permitted.[6]

We believe that the experimental method we report here, pioneered by G&J, is perfectly capable of testing these and additional hypotheses regarding the psychological manifestations of MSC's. All that is really necessary to make this or any experimental technique useful is to formulate our hypotheses in explicit enough fashion to render them testable.

ACKNOWLEDGEMENT

This research was supported in part by the National Science Foundation.

FOOTNOTES

1. The unabridged Oxford English Dictionary lists bleb, defining it as a smaller variety of blob. In all likelihood, though, this word is not in the lexicons of most speakers of English, and so this would not decrease the predicted difference between [klæb] and [klɛb].

2. A study by M. Ohala (1972) seems to provide additional indirect support for this. She asked Hindi-speaking subjects to categorize normal and rare Hindi words and made-up words as 'existing', 'possible but non-existing', and 'not possible' for Hindi. In some cases their anomalous assignment of what the linguist would consider 'accidental gaps' in the lexicon to the category 'not possible' would be accounted for by acknowledging that the 3-way classification was too restrictive and that they could have done better making scalar judgements.

3. G&J in their study used both an 11-point scale and the "free magnitude estimation" technique and found them both to give the same results. We used an 11-point scale since it was simpler to explain to subjects how to use it.

4. In the study by M. Ohala (1972) referred to in note 2, it was apparent that the subjects not only gave evidence of knowing the #SC's of Hindi but in some cases of knowing #SC's appropriate to different parts of the Hindi lexicon, e.g., what consonant clusters would be appropriate for English loanwords or words of Urdu origin.

5. In a conversation after this paper was given, Halle (personal communication) acknowledged that the proposals in SPE on the quantification of morpheme well-formedness were tenative and incorrect to the extent that they did not predict mløf to be further from English than spøf.

6. Bertz and Bever (1973), in their study of the awareness monolingual speakers of English have of universal consonant cluster patterns, accepted the G&J phoneme substitution measure with these arbitrary and unproven restrictions, i.e., that phonemes may not be added or deleted. They found that their subjects judged a word such as [ntif] to be more common in the world's languages than [nkif]. They concluded that since neither of these clusters occur in initial position in English and that since both words would be rated equally far from English using the G&J measure, their subjects couldn't have used their knowledge of English in making their judgement but must have used their knowledge of universal tendencies of consonant cluster formation. However, this conclusion, like ours, is only as good as the assumptions it is based on. If their subjects did use something like the G&J measure except that addition and deletion of phonemes were also permitted, then these initial clusters could become final clusters. In that case the homorganic nasal + stop cluster would certainly be judged preferrable or more common than the non-homorganic clusters and we need not assume the subjects had knowledge of sound patterns not in their native language.

REFERENCES

Chomsky, N. & Halle, M. (1968) The Sound Pattern of English. Harper and Row.

Esper, E.A. (1925) A technique for the experimental investigation of associative interference in artificial linguistic material. Language Monographs, No. 1.

Fromkin, V.A. (1971) The non-anomalous nature of anomalous utterances. Language 47:27-52.

Greenberg, J.H. & Jenkins, J.J. (1964) Studies in the psychological correlates of the sound system of American English. Word 20:157-177.

Halle, M. (1959) The Sound Pattern of Russian. Mouton.

Halle, M. (1962) Phonology in generative grammar. Word 18:54-72. Reprinted in: J.A. Fodor & J.J. Katz (eds), The Structure of Language. Readings in the Philosophy of Language. Englewood Cliffs, N.J.: Prentice-Hall Inc. Pp. 334-352.

Ohala, M. (1972) Topics in Hindi-Urdu Phonology. Unpublished doctoral dissertation, University of California, Los Angeles.

Pertz, D.L. & Bever, T.G. (1975) Sensitivity to phonological universals in children and adolescents. Language 51:149-162.

Stanley, R. (1967) Redundancy rules in phonology. Language 43:393-436

Wright, J. (1975) Nasal-stop assimilation: Testing the psychological reality of an English MSC. In: C.A. Ferguson, L.M. Hyman, and J.J. Ohala (eds.), Nasalfest: Papers from a Symposium on Nasals and Nasalization. Stanford: Language Universals Project. Pp. 389-397.

Zimmer, K. (1969) Psychological correlates of some Turkish morpheme structure conditions. Language 46:309-321.

ANNOUNCEMENT

Language Tapes Available

As a supplement to the introductory phonetics and phonology course offered by the Department of Linguistics, the Phonology Laboratory has made several language tapes designed to give students practice in discrimination of and IPA transcription of important sound contrasts. Copies of some of these tapes may be purchased from the Phonology Laboratory.

All the tapes have the same format:

Part I: Examples of words containing what are considered the "interesting" sounds of the language. Each word is preceded by a number (corresponding to its number on the accompanying text), is repeated twice in isolation and once in a simple sentence frame (which is identical for all sample words). E.g., "Number thirteen (pause) [atʰa] (pause) [atʰa] (pause) [vo atʰa hɛ]."

Part II: Words from Part I spoken in immediate juxtaposition to illustrate minimal pairs, triplets, or n-tuples. Each justaposed series is preceded by the numbers of the words, then are pronounced together in isolation twice, and finally once in the sentence frame. E.g., "Numbers thirteen and nineteen (pause) [atʰa] [atʰa] (pause) [atʰa][atʰa] (pause) [vo atʰa hɛ][vo atʰa hɛ]."

Part III: Words not included in Parts I and II but made up of sounds exemplified in those parts. These are included to give the listener a chance to practice IPA phonetic transcription. The format is otherwise similar to that used in the preceding parts.

A complete text accompanies each tape. This includes a) a brief description of the language and the dialect of the speaker, b) an explanation of any special phonetic symbols used (although, in general, IPA conventions are adhered to), c) the text of the tape, including the English translations of the words spoken, and d) the standard native ortographic transcription of all words on the tape.

At present the tapes are available only on a reel. If there is sufficient demand they could be provided in casette. All copies are made on archive-quality tape with a playback speed of 7.5 ips (19 cm/sec) and are 1/2 track.

At present we have available tapes for the following languages: French (Picardy dialect), Quechua, Hindi, Mandarin Chinese, and Thai. Tapes for other languages are in preparation. The cost of these tapes (until further notice) is $18.00 (U.S.) each, including surface post. Please allow 2 months for an order to be filled.